```
                     clone_end:T7
                     vector_side:right"
     misc_feature    144606. .187852
                     /note="assembly_name:Contig41"
SE COUNT     55964 a  35742 c  36408 g  58936 t    802 others
IGIN

Query Match             45.7%;  Score 738;  DB 2;  Length 187852;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  738;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      264 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    64776 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 64835

      324 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 383
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    64836 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 64895

      384 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    64896 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 64955

      444 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    64956 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 65015

      504 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65016 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 65075

      564 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65076 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 65135

      624 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65136 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 65195

      684 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65196 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 65255

      744 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65256 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 65315

      804 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65316 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 65375

      864 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65376 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 65435

      924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    65436 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 65495

      984 AGAGCCTCTTTGAAAGGT 1001
          ||||||||||||||||||
    65496 AGAGCCTCTTTGAAAGGT 65513


ULT 5
5s8380/c
US       AX458380                 621 bp    DNA     linear   PAT 08-JUL-2002
INITION  Sequence 5 from Patent WO0246411.
ESSION   AX458380
SION     AX458380.1  GI:21725048
WORDS    .
RCE      human.
RGANISM  Homo sapiens
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
  AUTHORS  Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE    Intestinal proteins
  JOURNAL  Patent: WO 0246411-A 5 13-JUN-2002;
           Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 71851705V1"
BASE COUNT      142 a    138 c    179 g    162 t
ORIGIN

  Query Match             34.5%;  Score 558;  DB 6;  Length 621;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-305;
  Matches  608;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    210 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 562

Qy    270 GGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGTCCTTAT 329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 GGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGTCCTTAT 502

Qy    330 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 442

Qy    390 AAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGAACACCA 449
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    441 AAATCATAGAGAAACTAGATCAGCAGACCCCACCCAGACCTTTCACCCACGTGAACACCA 382

Qy    450 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 322

Qy    510 GAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 GAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 262

Qy    570 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 202

Qy    630 AGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 AGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 142

Qy    690 AGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTG 749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 AGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTG 82

Qy    750 CAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCC 809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 CAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCC 22

Qy    810 AAGTCCACT 818
          |||||||||
Db     21 AAGTCCACT 13


RESULT 10
AK000134
LOCUS       AK000134                1704 bp    mRNA    linear   PRI 22-FEB-2000
DEFINITION  Homo sapiens cDNA FLJ20127 fis, clone COL06176.
ACCESSION   AK000134
VERSION     AK000134.1  GI:7020022
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL06176.
  ORGANISM  Homo sapiens
```

```
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE   1  (sites)
AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
          Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
          Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE     NEDO human cDNA sequencing project
JOURNAL   Unpublished
FERENCE   2  (bases 1 to 1704)
AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
          Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE     Direct Submission
JOURNAL   Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1,
          Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp,
          Tel:81-3-5449-5286, Fax:81-3-5449-5416)
MMENT     NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing: Research Association for Biotechnology; cDNA library
          construction, 5'- & 3'-end one pass sequencing: Departent of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
ATURES              Location/Qualifiers
   source           1. .1704
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="COL06176"
                    /tissue_type="colon"
                    /clone_lib="COL"
                    /note="cloning vector pME18SFL3"
   CDS              558. .1040
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="BAA90965.1"
                    /db_xref="GI:7020023"
                    /translation="MEKFNTISASKCNKETVAMKEKCKFGMTSTIPSGHVWRNTWNPV
                    SCSLATVKMKECLRGKLIYLMGDSTIRQWMEYFKASINTLKSVDLHESGKLQHQLAVD
                    LDRNINIQWQKYCYPLIGSMTYSVKEMEYLTRAIDRTGGEKKYCHCYFPGPAFQTLSH
                    "
SE COUNT      597 a    366 c    308 g    433 t
IGIN

Query Match           32.7%;  Score 528;  DB 9;  Length 1704;
Best Local Similarity  100.0%;  Pred. No. 1.7e-288;
Matches  528;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTTGGTATAGACAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTTGGTATAGACAA 60

     61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

    121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

    181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240

    241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300

    301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360

    361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420

Qy    421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480

Qy    481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCT 528


RESULT 11
AX458381/c
LOCUS       AX458381                545 bp    DNA     linear   PAT 08-JUL-2002
DEFINITION  Sequence 6 from Patent WO0246411.
ACCESSION   AX458381
VERSION     AX458381.1  GI:21725049
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1.
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 6 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 70255975V1"
BASE COUNT      110 a    129 c    161 g    145 t
ORIGIN

  Query Match           25.3%;  Score 409;  DB 6;  Length 545;
  Best Local Similarity   99.6%;  Pred. No. 9.9e-221;
  Matches  509;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    319 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 452

Qy    379 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCGATCACCCA 392

Qy    439 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 332

Qy    499 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 272

Qy    559 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 212

Qy    619 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 152

Qy    679 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 92

Qy    739 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     91 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 32

Qy    799 TGGCACTTCCCAAGTCCACTCTGAATGTGGC 829
          |||||||||||||||||||||||||||||||
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:33:06 ; Search time 2189 Seconds
(without alignments)
11956.102 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence: 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	268	16.6	358	9	AI833131	AI833131 at75d06.x
	2	169	10.5	905	12	HG199761	BG199761 RST19055
	3	148	9.2	343	12	BG201797	BG201797 RST21139
c	4	112	6.9	112	9	AI581835	AI581835 ar94g08.x
	5	111	6.9	946	12	BG217603	BG217603 RST37315
c	6	109	6.7	530	17	AQ439960	AQ439960 HS_5059_B
	7	88	5.4	634	13	BI764226	BI764226 603045907
c	8	45	2.8	327	10	AW842215	AW842215 RC0-CN002
	9	32	2.0	664	10	BB625007	BB625007 BB625007
	10	32	2.0	3267	11	AK019469	AK019469 Mus muscu
	11	27	1.7	573	17	AZ454451	AZ454451 1M0256J19
	12	27	1.7	630	10	BB656245	BB656245 BB656245
	13	24	1.5	378	12	BF721682	BF721682 mab21c06.
	14	24	1.5	446	10	BB862365	BB862365 BB862365
	15	24	1.5	597	12	BG088025	BG088025 H3147F12-
	16	24	1.5	671	10	BB605016	BB605016 BB605016
	17	23	1.4	208	14	R75193	R75193 MDB1137 Mou
c	18	23	1.4	227	13	BG942595	BG942595 ax27b03.x
c	19	23	1.4	252	10	AW522484	AW522484 UI-R-B00-
c	20	23	1.4	279	10	AW087776	AW087776 xb69c11.x
c	21	23	1.4	357	9	AI136832	AI136832 UI-R-C2p-
c	22	23	1.4	375	12	BG514625	BG514625 dae09f12.
c	23	23	1.4	403	9	AI482526	AI482526 vg49d02.x
c	24	23	1.4	418	10	AW494429	AW494429 UI-M-BH3-
c	25	23	1.4	418	13	BI884210	BI884210 fo20a06.x
c	26	23	1.4	434	13	BI494352	BI494352 df109e02.
	27	23	1.4	436	9	AL047005	AL047005 DKFZp586M
c	28	23	1.4	472	10	AW435156	AW435156 UI-R-BJ0p
	29	23	1.4	490	13	BM445719	BM445719 1IL14H4.a
c	30	23	1.4	492	14	BQ018316	BQ018316 UI-H-DH1-
	31	23	1.4	495	14	C93956	C93956 C93956 Dict
c	32	23	1.4	521	12	BG579297	BG579297 dab85b07.
c	33	23	1.4	536	12	BF510845	BF510845 UI-H-BI4-
c	34	23	1.4	537	13	BM385280	BM385280 UI-R-EB1-
c	35	23	1.4	544	14	BM989987	BM989987 UI-H-DI0-
c	36	23	1.4	563	10	AW526128	AW526128 UI-R-B01-
c	37	23	1.4	564	13	BI402698	BI402698 MI-P-CP0-
	38	23	1.4	580	13	BM030140	BM030140 488665 MA
c	39	23	1.4	588	14	BM973554	BM973554 UI-CF-EC1
c	40	23	1.4	598	17	BH323298	BH323298 CH230-20N
c	41	23	1.4	620	14	BQ574501	BQ574501 UI-H-EZ1-
	42	23	1.4	622	9	AU039013	AU039013 AU039013
c	43	23	1.4	622	14	BM683810	BM683810 UI-E-EJ1-
c	44	23	1.4	623	14	BM971486	BM971486 UI-CF-EC1
c	45	23	1.4	639	14	BQ190688	BQ190688 UI-R-DN1-

ALIGNMENTS

RESULT 1
AI833131/c
LOCUS AI833131 358 bp mRNA linear EST 13-JUL-1999
DEFINITION at75d06.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2377835 3' similar to SW:BB61_RABIT Q05004 BRUSH BORDER 61.9 KD PROTEIN PRECURSOR. ;, mRNA sequence.
ACCESSION AI833131
VERSION AI833131.1 GI:5455111
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

```
FEATURES             Location/Qualifiers
     source          1. .358
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2377835"
                     /clone_lib="Barstead colon HPLRB7"
                     /sex="male"
                     /dev_stage="adult, age 25"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT      106 a     60 c     74 g    118 t
ORIGIN

  Query Match             16.6%;  Score 268;  DB 9;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-76;
  Matches  268;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1328 AAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAA 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 AAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAA 209

Qy   1388 AACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAA 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 AACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAA 149

Qy   1448 TATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGAT 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 TATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGAT 89

Qy   1508 ATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAG 1567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 ATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAG 29

Qy   1568 ATTAATATATTATTAAACTATATTTGTT 1595
          |||||||||||||||||||||||||||
Db     28 ATTAATATATTATTAAACTATATTTGTT 1


RESULT 2
BG199761
LOCUS       BG199761                905 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST19055 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG199761
VERSION     BG199761.1  GI:13721448
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 387.
FEATURES             Location/Qualifiers
     source          1. .905
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      228 a    226 c    296 g    154 t      1 others
ORIGIN

  Query Match             10.5%;  Score 169;  DB 12;  Length 905;
  Best Local Similarity   99.1%;  Pred. No. 6e-45;
  Matches  319;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    352 ATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCA 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 ATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCA 132

Qy    412 GCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCAC 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 GCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCAC 192

Qy    472 AGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCT 531
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    193 AGCCACCATCCTCAACCCTCGAGATACGTAGTGGAGGGGAGACCAGCTGCACATCCTGCT 252

Qy    532 GGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAG 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 GGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAG 312

Qy    592 GATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGG 651
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    313 GATGTCTTCCCCAGCGCTGATGGGAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGG 372

Qy    652 CACCTACCTGGTCAGCTTCACT 673
          |||||||||||||||||||||||
Db    373 CACCTACCTGGTCAGCTTCACT 394


RESULT 3
BG201797
LOCUS       BG201797                343 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST21139 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG201797
VERSION     BG201797.1  GI:13723484
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 293.
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:33:06 ; Search time 2189 Seconds
(without alignments)
11956.102 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence: 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	268	16.6	358	9	AI833131	AI833131 at75d06.x
	2	169	10.5	905	12	BG199761	BG199761 RST19055
	3	148	9.2	343	12	BG201797	BG201797 RST21139
c	4	112	6.9	112	9	AI581835	AI581835 ar94g08.x
	5	111	6.9	946	12	BG217603	BG217603 RST37315
c	6	109	6.7	530	17	AQ439960	AQ439960 HS_5059_B
	7	88	5.4	634	13	BI764226	BI764226 603045907
c	8	45	2.8	327	10	AW842215	AW842215 RC0-CN002
	9	32	2.0	664	10	BB625007	BB625007 BB625007
	10	32	2.0	3267	11	AK019469	AK019469 Mus muscu
	11	27	1.7	573	17	AZ454451	AZ454451 1M0256J19
	12	27	1.7	630	10	BB656245	BB656245 BB656245
	13	24	1.5	378	12	BF721682	BF721682 mab21c06.
	14	24	1.5	446	10	BB862365	BB862365 BB862365
	15	24	1.5	597	12	BG088025	BG088025 H3147F12-
	16	24	1.5	671	10	BB605016	BB605016 BB605016
	17	23	1.4	208	14	R75193	R75193 MDB1137 Mou
c	18	23	1.4	227	13	BG942595	BG942595 ax27b03.x
c	19	23	1.4	252	10	AW522484	AW522484 UI-R-BO0-
c	20	23	1.4	279	10	AW087776	AW087776 xb69c11.x
c	21	23	1.4	357	9	AI136832	AI136832 UI-R-C2p-
c	22	23	1.4	375	12	BG514625	BG514625 dae09f12.
c	23	23	1.4	403	9	AI482526	AI482526 vg49d02.x
c	24	23	1.4	418	10	AW494429	AW494429 UI-M-BH3-
c	25	23	1.4	418	13	BI884210	BI884210 fo20a06.x
c	26	23	1.4	434	13	BI494352	BI494352 df109e02.
	27	23	1.4	436	9	AL047005	AL047005 DKFZp586M
c	28	23	1.4	472	10	AW435156	AW435156 UI-R-BJ0p
	29	23	1.4	490	13	BM445719	BM445719 1IL14H4.a
c	30	23	1.4	492	14	BQ018316	BQ018316 UI-H-DH1-
	31	23	1.4	495	14	C93956	C93956 C93956 Dict
c	32	23	1.4	521	12	BG579297	BG579297 dab85b07.
c	33	23	1.4	536	12	BF510845	BF510845 UI-H-BI4-
c	34	23	1.4	537	13	BM385280	BM385280 UI-R-EB1-
c	35	23	1.4	544	14	BM989987	BM989987 UI-H-DI0-
c	36	23	1.4	563	10	AW526128	AW526128 UI-R-BO1-
c	37	23	1.4	564	13	BI402698	BI402698 MI-P-CP0-
	38	23	1.4	580	13	BM030140	BM030140 488665 MA
c	39	23	1.4	588	14	BM973554	BM973554 UI-CF-EC1
c	40	23	1.4	598	17	BH323298	BH323298 CH230-20N
c	41	23	1.4	620	14	BQ574501	BQ574501 UI-H-EZ1-
	42	23	1.4	622	9	AU039013	AU039013 AU039013
c	43	23	1.4	622	14	BM683810	BM683810 UI-E-EJ1-
c	44	23	1.4	623	14	BM971486	BM971486 UI-CF-EC1
c	45	23	1.4	639	14	BQ190688	BQ190688 UI-R-DN1-

ALIGNMENTS

RESULT 1
AI833131/c
LOCUS AI833131 358 bp mRNA linear EST 13-JUL-1999
DEFINITION at75d06.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2377835 3' similar to SW:BB61_RABIT Q05004 BRUSH BORDER 61.9 KD PROTEIN PRECURSOR. ;, mRNA sequence.
ACCESSION AI833131
VERSION AI833131.1 GI:5455111
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

```
FEATURES             Location/Qualifiers
     source          1. .358
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2377835"
                     /clone_lib="Barstead colon HPLRB7"
                     /sex="male"
                     /dev_stage="adult, age 25"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT       106 a     60 c     74 g    118 t
ORIGIN

  Query Match             16.6%;  Score 268;  DB 9;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-76;
  Matches  268;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1328 AAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAA 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 AAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAA 209

Qy   1388 AACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAA 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 AACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAA 149

Qy   1448 TATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGAT 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 TATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGAT 89

Qy   1508 ATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAG 1567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 ATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAG 29

Qy   1568 ATTAATATATTATTAAACTATATTTGTT 1595
          |||||||||||||||||||||||||||
Db     28 ATTAATATATTATTAAACTATATTTGTT 1

RESULT 2
BG199761
LOCUS       BG199761                905 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST19055 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG199761
VERSION     BG199761.1  GI:13721448
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Cranlata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhinl; Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide proteln expression llbraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 387.
FEATURES             Location/Qualifiers
     source          1. .905
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_llb="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT       228 a    226 c    296 g    154 t      1 others
ORIGIN

  Query Match             10.5%;  Score 169;  DB 12;  Length 905;
  Best Local Similarity   99.1%;  Pred. No. 6e-45;
  Matches  319;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    352 ATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCA 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 ATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCA 132

Qy    412 GCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCAC 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 GCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCAC 192

Qy    472 AGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 AGCCACCATCCTCAACCCTCGAGATACGTAGTGGAGGGGAGACCAGCTGCACATCCTGCT 252

Qy    532 GGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAG 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 GGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAG 312

Qy    592 GATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGATTTCAACAACGG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 GATGTCTTCCCCAGCGCTGATGGGAGGTGCTTCAGGAAAGGTGACTGATTTCAACAACGG 372

Qy    652 CACCTACCTGGTCAGCTTCACT 673
          |||||||||||||||||||||||
Db    373 CACCTACCTGGTCAGCTTCACT 394

RESULT 3
BG201797
LOCUS       BG201797                343 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST21139 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence
ACCESSION   BG201797
VERSION     BG201797.1  GI:13723484
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 293.
```

```
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT       102 a    100 c     85 g     56 t
ORIGIN

  Query Match             9.2%;  Score 148;  DB 12;  Length 343;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-38;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      355 AAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCA 414
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       81 AAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCA 140

Qy      415 GATCCCACCCAGACCTTTCACCCACGTCAACACCACCACCAGCGCCACACATAGCACAGC 474
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      141 GATCCCACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGC 200

Qy      475 CACCATCCTCAACCCTCGAGATACGTAC 502
            ||||||||||||||||||||||||||||
Db      201 CACCATCCTCAACCCTCGAGATACGTAC 228


RESULT 4
AI581835/c
LOCUS        AI581835                  112 bp    mRNA    linear   EST 06-APR-1999
DEFINITION   ar94q08.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
             IMAGE:2173310 3' similar to SW:BR61_RABIT Q05004 BRUSH BORDER 61.9
             KD PROTEIN PRECURSOR. ;, mRNA sequence.
ACCESSION    AI581835
VERSION      AI581835.1  GI:4567732
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 112)
  AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
             Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
             ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
             White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE      WashU-NCI human EST Project
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .112
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2173310"
                     /clone_lib="Barstead colon HPLRB7"
                     /sex="male"
                     /dev_stage="adult, age 25"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT       36 a     18 c     18 g     40 t
ORIGIN

  Query Match             6.9%;  Score 112;  DB 9;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-26;
  Matches  112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1484 GTGAGTATCATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCA 1543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      112 GTGAGTATCATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCA 53

Qy     1544 CCTCAACATGTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTT 1595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       52 CCTCAACATGTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTT 1


RESULT 5
BG217603
LOCUS        BG217603                  946 bp    mRNA    linear   EST 21-APR-2001
DEFINITION   RST37315 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION    BG217603
VERSION      BG217603.1  GI:13743624
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 946)
  AUTHORS    Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
             Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
             , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
             ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
             ,J., Danzig,J. and Ducar,M.
  TITLE      Creation of genome-wide protein expression libraries using random
             activation of gene expression
  JOURNAL    Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE    21227151
COMMENT      Contact: Scott J. Cain
             Athersys, Inc.
             3201 Carnegie Ave, Cleveland, OH 44115, USA
             Tel: 216 431 9900
             Fax: 216 361 9596
             Email: scain@athersys.com
             High quality sequence stop: 323.
FEATURES             Location/Qualifiers
     source          1. .946
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      225 a    221 c    329 g    169 t      2 others
ORIGIN

  Query Match             6.9%;  Score 111;  DB 12;  Length 946;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-26;
  Matches  111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      392 ATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACC 451
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      134 ATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGAACACCACC 193
```

```
Qy     452 ACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTAC 502
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 ACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTAC 244


RESULT 6
AQ439960/c
LOCUS       AQ439960                 530 bp    DNA     linear   GSS 31-MAR-1999
DEFINITION  HS_5059_B2_F05_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=635 Col=10 Row=L, DNA sequence.
ACCESSION   AQ439960
VERSION     AQ439960.1  GI:4551299
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 635  row: L  column: 10
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 530.
FEATURES                Location/Qualifiers
     source             1. .530
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="Plate=635 Col=10 Row=L"
                        /clone_lib="RPCI-11 Human Male BAC Library"
                        /sex="male"
                        /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                        Male blood DNA was isolated from one randomly chosen donor
                        and partially digested with a combination of EcoRI and
                        EcoRI Methylase. Size selected DNA was cloned into the
                        pBACe3.6 vector at EcoRI sites"
BASE COUNT      143 a    130 c     86 g    162 t      9 others
ORIGIN

  Query Match             6.7%;  Score 109;  DB 17;  Length 530;
  Best Local Similarity   99.4%;  Pred. No. 1.9e-25;
  Matches  159;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     398 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 339

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     338 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACTTAGCTCGCCAG 279

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGG 160
           ||||||||||||||||||||||||||||||||||||||||
Db     278 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGG 239


RESULT 7
BI764226
LOCUS       BI764226                 634 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603045907F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186472 5',
            mRNA sequence.
ACCESSION   BI764226
VERSION     BI764226.1  GI:15755804
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11465  row: m  column: 01
            High quality sequence stop: 634.
FEATURES                Location/Qualifiers
     source             1. .634
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5186472"
                        /clone_lib="NIH_MGC_116"
                        /lab_host="DH10B"
                        /note="Organ: pooled colon, kidney, stomach; Vector:
                        pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                        source anonymous pool of 3 colons, age 26 yo male, 49 yo
                        female, 71 yo male colon; 46 yo male kidney, and pool of 2
                        stomachs, 62 yo male and 70 yo female.  Library is
                        oligo-dT primed and  directionally cloned (EcoRV site is
                        destroyed upon cloning).  Average insert size 1.4 kb,
                        insert size range 1-3 kb.  Library is normalized and
                        enriched for full-length clones and was constructed by C.
                        Gruber  (Invitrogen).  Research Genetics tracking code
                        023. Note: this is a NIH_MGC Library."
BASE COUNT      180 a    175 c    140 g    139 t
ORIGIN

  Query Match             5.4%;  Score 88;  DB 13;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-19;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 234

Qy     411 AGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||
Db     235 AGCAGATCCCACCCAGACCTTTCACCCA 262


RESULT 8
AW842215/c
LOCUS       AW842215                 327 bp    mRNA    linear   EST 18-MAY-2000
DEFINITION  RC0-CN0028-010200-012-g11 CN0028 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW842215
VERSION     AW842215.1  GI:7936198
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 327)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC0-CN0028-010
            200-012-g11&t3=2000-02-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 39
            High quality sequence stop: 327.
FEATURES             Location/Qualifiers
     source          1. .327
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CN0028"
                     /dev_stage="Adult"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       77 a    104 c     48 g     98 t
ORIGIN

  Query Match             2.8%;  Score 45;  DB 10;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 0.00011;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     116 GCCAGAGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGG 160
           |||||||||||||||||||||||||||||||||||||||||||||
Db     311 GCCAGAGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGG 267


RESULT 9
BH625007
LOCUS       BB625007                 664 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB625007 RIKEN full-length enriched, adult male colon Mus musculus
            cDNA clone 9030614N15 5', mRNA sequence.
ACCESSION   BB625007
VERSION     BB625007.1  GI:16463365
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="9030614N15"
                     /clone_lib="RIKEN full-length enriched, adult male colon"
                     /sex="male"
                     /tissue_type="colon"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 185.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:
                     BamHI"
BASE COUNT      161 a    183 c    165 g    155 t
ORIGIN

  Query Match             2.0%;  Score 32;  DB 10;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     716 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAG 747
           ||||||||||||||||||||||||||||||||
Db     524 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAG 555
```

RESULT 10
AK019469
LOCUS AK019469 3267 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus 14 days embryo liver cDNA, RIKEN full-length enriched library, clone:4432416J03:homolog to BRUSH BORDER 61.9 KDA PROTEIN PRECURSOR, full insert sequence.
ACCESSION AK019469
VERSION AK019469.1 GI:12859693
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 14 days embryo liver cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:4432416J03.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5 (bases 1 to 3267)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC) RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES Location/Qualifiers
source 1. .3267
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4432416J03"
/db_xref="MGD:MGI:1912173"
/db_xref="taxon:10090"
/clone="4432416J03"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14 days embryo"
CDS 205. .1881
/note="data source:SPTR, source key:Q05604, evidence:ISS homolog to BRUSH BORDER 61.9 KDA PROTEIN PRECURSOR putative"
/codon_start=1
/protein_id="BAB31740.1"
/db_xref="GI:12859694"
/db_xref="MGD:MGI:1925502"
/translation="MRRMLSPRILLSSLPNASAQKLFLIVLIIFAFWVVFMTSKDHTE
FMVHLNNRIILRRWSIFKEFLHSEELKNTPASVEAELAVTAILEKLNQQIPPRPFQTH
SSTTSAKQSTATIHNPQRTYCVGDQLNVLLVAKDYFGNRKEYGGDFLRARMSSPALKA
GASGKVTDFNNGTYLISFTLFWEGPVSLSILLMHPSEGVSALWRARNRGYGKIIFIAC
FLNGTSPVLTECGLTLNTSAELCQYLDARDHEAFYCLKLPGVPCEALTHMTSKNSNIS
YLSLEEKLLFRRFNIGVEVVKNLSIVVSPCNNKTNKKKKKCQIGMETPSPGGYTLKQH
WITAHCEQNEFRAIKDINNCLTRKLIYLMGDSTLRQWIYYLPKVVKTLKFFDDHGAGP
FKTHILLDTERHIFVQWKKHSHPFVTNKLFSMKDDNYIPREIDQVAGDSGTAIVISFG
QHFRPFPINVFIRRAINVKNAIERLFLRSPETKVIIKTENIREINEHVFIPSDFHGSI
QNLIIRDIFRDLNVGIIDAWDMTVAYRSEDVHPPESVIESQIGMFLNYIC"
polyA_signal 3250. .3255
/note="putative"
polyA_site 3267
/note="putative"
BASE COUNT 954 a 726 c 673 g 914 t
ORIGIN

Query Match 2.0%; Score 32; DB 11; Length 3267;

```
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     716 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAG 747
           ||||||||||||||||||||||||||||||||
Db     784 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAG 815


RESULT 11
AZ454451
LOCUS       AZ454451                573 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0256J19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0256J19 F, DNA sequence.
ACCESSION   AZ454451
VERSION     AZ454451.1  GI:10612576
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0256  row: J  column: 19
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 573.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0256J19"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      152 a    161 c    125 g    135 t
ORIGIN

  Query Match             1.7%;  Score 27;  DB 17;  Length 573;
```

```
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     662 GTCAGCTTCACTCTGTTCTGGGAGGGC 688
           |||||||||||||||||||||||||||
Db     511 GTCAGCTTCACTCTGTTCTGGGAGGGC 537


RESULT 12
BB656245
LOCUS       BB656245                630 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB656245 RIKEN full-length enriched, 12 days embryo spinal ganglion
            Mus musculus cDNA clone D130007D15 5', mRNA sequence.
ACCESSION   BB656245
VERSION     BB656245.1  GI:16490073
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="D130007D15"
                     /clone_lib="RIKEN full-length enriched, 12 days embryo
                     spinal ganglion"
                     /tissue_type="spinal ganglion"
                     /dev_stage="12 days embryo"
```

/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 174 a 167 c 146 g 143 t
ORIGIN

Query Match 1.7%; Score 27; DB 10; Length 630;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 TCAGCTTCACTCTGTTCTGGGAGGGCC 689
||||||||||||||||||||||||||||
Db 595 TCAGCTTCACTCTGTTCTGGGAGGGCC 621

RESULT 13
BF721682
LOCUS BF721682 378 bp mRNA linear EST 03-JAN-2001
DEFINITION mab21c06.x1 Soares_NMEBA_branchial_arch Mus musculus cDNA clone IMAGE:3970931 3' similar to SW:BB61_RABIT Q05004 BRUSH BORDER 61.9 KD PROTEIN PRECURSOR. ;, mRNA sequence.
ACCESSION BF721682
VERSION BF721682.1 GI:12022684
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 378)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/iresources.shtml

MGI:1470963
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 241.
FEATURES Location/Qualifiers
source 1. .378
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3970931"
/clone_lib="Soares_NMEBA_branchial_arch"
/tissue_type="branchial arches"
/dev_stage="embryo, 10.5 dpc"
/lab_host="DH10B (phage resistant)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATGCATTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 106 a 101 c 84 g 87 t
ORIGIN

Query Match 1.5%; Score 24; DB 12; Length 378;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 410 CAGCAGATCCCACCCAGACCTTTC 433
||||||||||||||||||||||||
Db 190 CAGCAGATCCCACCCAGACCTTTC 213

RESULT 14
BB862365
LOCUS BB862365 446 bp mRNA linear EST 26-NOV-2001
DEFINITION BB862365 RIKEN full-length enriched, pooled cell lines Mus musculus cDNA clone G430027D22 5', mRNA sequence.
ACCESSION BB862365
VERSION BB862365.1 GI:17103819
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 446)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,I., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Taqawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al., 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
Iwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,I., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .446
/organism="Mus musculus"
/db_xref="taxon:10090"

```
                     /clone="G430027D22"
                     /clone_lib="RIKEN full-length enriched, pooled cell lines"
                     /note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164),
                     (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
                     (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
                     (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16
                     melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
                     ), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
                     (cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
                     (tissue_type=bladder, cell_line=RCB-0544 MBT-2),
                     (tissue_type=bone marrow, cell_type=stroma cell,
                     cell_line= CRL-2028 SR-4987), (tissue_type=colon,
                     cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
                     cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
                     cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
                     cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
                     ), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
                     BC3H1)"
BASE COUNT      110 a    134 c    110 g     92 t
ORIGIN

  Query Match             1.5%;  Score 24;  DB 10;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+02;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     410 CAGCAGATCCCACCCAGACCTTTC 433
           ||||||||||||||||||||||||
Db      12 CAGCAGATCCCACCCAGACCTTTC 35


RESULT 15
BG088025
LOCUS       BG088025                597 bp    mRNA    linear   EST 26-JAN-2001
DEFINITION  H3147F12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3147F12 5', mRNA sequence.
ACCESSION   BG088025
VERSION     BG088025.1  GI:12570589
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3147F12-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3147  row: F  column: 12
            Seq primer: -21M13 Reverse
            High quality sequence stop: 597
            POLYA-No.
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3147F12-5"
                     /db_xref="taxon:10090"
                     /clone="H3147F12"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo cDNA libraries (including preimplantation stage
                     embryos from unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos, extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad) and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers. References include: (1) Genome-wide
                     expression profiling of mid-gestation placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray, 2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA analysis reveals phased gene expression
                     patterns during preimplantation mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day mouse embryos reveals enrichment in the t-complex
                     and under-representation on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT      169 a    151 c    138 g    139 t
ORIGIN

  Query Match             1.5%;  Score 24;  DB 12;  Length 597;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     410 CAGCAGATCCCACCCAGACCTTTC 433
           ||||||||||||||||||||||||
Db     449 CAGCAGATCCCACCCAGACCTTTC 472


Search completed: November 22, 2002, 04:29:06
Job time : 2198 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:38:36 ; Search time 866 Seconds
(without alignments)
706.730 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence: 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 335578 seqs, 189365133 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1616	100.0	1616	10	US-09-729-454-3	Sequence 3, Appli
	2	1005	62.2	1005	10	US-09-729-454-30	Sequence 30, Appl
	3	560	34.7	1212	9	US-09-981-353-144	Sequence 144, App
c	4	558	34.5	621	10	US-09-729-454-5	Sequence 5, Appli
	5	488	30.2	841	9	US-09-981-353-145	Sequence 145, App
c	6	409	25.3	545	10	US-09-729-454-6	Sequence 6, Appli
	7	375	23.2	394	10	US-09-729-454-9	Sequence 9, Appli
	8	240	14.9	240	10	US-09-729-454-4	Sequence 4, Appli
	9	213	13.2	236	10	US-09-729-454-7	Sequence 7, Appli
	10	198	12.3	554	9	US-10-046-935-181	Sequence 181, App
c	11	93	5.8	580	9	US-10-046-935-388	Sequence 388, App
	12	88	5.4	661	10	US-09-729-454-11	Sequence 11, Appl
	13	88	5.4	1545	10	US-09-729-454-31	Sequence 31, Appl
	14	88	5.4	2248	10	US-09-729-454-10	Sequence 10, Appl
	15	73	4.5	414	10	US-09-729-454-8	Sequence 8, Appli
c	16	47	2.9	518	10	US-09-729-454-12	Sequence 12, Appl
	17	34	2.1	244	10	US-09-729-454-18	Sequence 18, Appl
	18	34	2.1	270	10	US-09-729-454-19	Sequence 19, Appl
c	19	34	2.1	291	10	US-09-729-454-17	Sequence 17, Appl
	20	32	2.0	206	10	US-09-729-454-16	Sequence 16, Appl
	21	32	2.0	257	10	US-09-729-454-22	Sequence 22, Appl
	22	25	1.5	288	10	US-09-729-454-20	Sequence 20, Appl
	23	24	1.5	275	10	US-09-729-454-21	Sequence 21, Appl
c	24	24	1.5	394	10	US-09-729-454-9	Sequence 9, Appli
	25	22	1.4	91	10	US-09-878-574-5331	Sequence 5331, Ap
	26	22	1.4	144	10	US-09-729-454-27	Sequence 27, Appl
	27	22	1.4	250	10	US-09-729-454-24	Sequence 24, Appl
c	28	22	1.4	357	10	US-09-960-352-8389	Sequence 8389, Ap
c	29	22	1.4	396	10	US-09-825-294-40	Sequence 40, Appl
	30	22	1.4	4158	10	US-09-349-015-34	Sequence 34, Appl
	31	21	1.3	337	10	US-09-920-300A-507	Sequence 507, App
	32	21	1.3	337	12	US-10-033-528-507	Sequence 507, App
	33	21	1.3	365	10	US-09-783-590-2675	Sequence 2675, Ap
	34	21	1.3	380	9	US-10-079-623-188	Sequence 188, App
	35	21	1.3	381	9	US-10-079-623-36	Sequence 36, Appl
c	36	21	1.3	462	10	US-09-729-454-13	Sequence 13, Appl
c	37	21	1.3	531	10	US-09-729-454-14	Sequence 14, Appl
c	38	21	1.3	802	10	US-09-833-381-1895	Sequence 1895, Ap
	39	21	1.3	2456	10	US-09-965-313-3	Sequence 3, Appli
c	40	20	1.2	262	10	US-09-729-454-28	Sequence 28, Appl
	41	20	1.2	276	10	US-09-729-454-23	Sequence 23, Appl
c	42	20	1.2	625	10	US-09-770-149-750	Sequence 750, App
c	43	20	1.2	864	10	US-09-770-445-581	Sequence 581, App
	44	20	1.2	1448	10	US-09-764-853-153	Sequence 153, App
	45	20	1.2	1967	10	US-09-887-586A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-729-454-3
; Sequence 3, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1616
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 3229449CB1
US-09-729-454-3

Query Match 100.0%; Score 1616; DB 10; Length 1616;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
```

```
Qy    241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300

Qy    301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360

Qy    361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420

Qy    421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480

Qy    481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAG 540

Qy    541 GGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTC 600

Qy    601 CCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCT 660

Qy    661 GGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCC 720

Qy    721 CAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTT 780

Qy    781 CACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAA 840

Qy    841 CACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGT 900

Qy    901 GAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGT 960

Qy    961 TTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGAT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGAT 1020

Qy   1021 TATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTGGATCT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTGGATCT 1080

Qy   1081 GCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATCAACAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATCAACAT 1140

Qy   1141 CCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAAGAGAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAAGAGAT 1200

Qy   1201 GGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAAATACTGTCATTGTTAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTAT 1260

Qy   1261 TTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAA 1320

Qy   1321 TCTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAA 1380

Qy   1381 AACAGAAAACATCAGGCAGATGTACAATCATCCAGAAACATTTAGTGACTTTCATGGTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTA 1440

Qy   1441 CATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCACTGTGAGTATCATTGATGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCACTGTGAGTATCATTGATGC 1500

Qy   1501 CTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGGCATATAACAATTGCATATGGCACAAATAATCTACACCCACCTCAACATGTAGTCGG 1560

Qy   1561 AAATCAGATTAATATATTATTAAACTATATTTGTTAAATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAATCAGATTAATATATTATTAAACTATATTTGTTAAATAACAAAAAAAAAAAAAA 1616
```

RESULT 2
US-09-729-454-30
; Sequence 30, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 GNN.g9965027_000007_006
US-09-729-454-30

Query Match 62.2%; Score 1005; DB 10; Length 1005;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Caps 0.

```
Qy    593 ATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGC 652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGCAAAGGTGACTGACTTCAACAACGGC 60

Qy    653 ACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTCCTGCTC 712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTCCTGCTC 120

Qy    713 ATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGCAGTGCAAGGAACCAACGCTATCACAGG 772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAACGAACCAAGGCTATCACAGG 180

Qy    773 GTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTG 832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTCAATGTGGCCTG 240

Qy    833 ATCCTAAACACAAATGCTGAATTGTGCCACTACCTGGACAACACAGACCAAGAAGGCTTC 892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTC 300

Qy    893 TACTGTGTGAGGCCTCAACACATGCCCTGTGCTCCACTCACTCACATGTATTCTAAGAAC 952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TACTGTGTGAGGCCTCAACACATGCCCTGTGCTCCACTCACTCACATGTATTCTAAGAAC 360

Qy    953 AAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGT 1012
```

```
Db     361 AAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGT 420

Qy    1013 GTAGAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCA 1072
Db     421 GTAGAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCA 480

Qy    1073 GTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAAC 1132
Db     481 GTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAAC 540

Qy    1133 ATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTC 1192
Db     541 ATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTC 600

Qy    1193 AAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTC 1252
Db     601 AAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTC 660

Qy    1253 ATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGG 1312
Db     661 ATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGG 720

Qy    1313 GCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTT 1372
Db     721 GCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTT 780

Qy    1373 ATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTT 1432
Db     781 ATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTT 840

Qy    1433 CATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATC 1492
Db     841 CATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATC 900

Qy    1493 ATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACAT 1552
Db     901 ATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACAT 960

Qy    1553 GTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTTAA 1597
Db     961 GTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTTAA 1005


RESULT 3
US-09-981-353-144
; Sequence 144, Application US/09981353
  Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
   FILE REFERENCE: PA-0038 US
   CURRENT APPLICATION NUMBER: US/09/981,353
   CURRENT FILING DATE:  2001-10-11
   NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 144
;   LENGTH: 1212
    TYPE: DNA
    ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 1397029.1
US-09-981-353-144

  Query Match             34.7%;  Score 560;  DB 9;  Length 1212;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-232;
  Matches  730;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACT-ATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTT 239
Db     181 TATGATAAATTATAAGTCCCTAATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTT 240

Qy     240 TTACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCC 299
Db     241 TTACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCC 300

Qy     300 ATTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGC 359
Db     301 ATTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGC 360

Qy     360 CACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCC 419
Db     361 CACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCC 420

Qy     420 CACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCA 479
Db     421 CACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCA 480

Qy     480 TCCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGA 539
Db     481 TCCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGA 540

Qy     540 GGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTT 599
Db     541 GGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTT 600

Qy     600 CCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACC 659
Db     601 CCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACC 660

Qy     660 TGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACC 719
Db     661 TGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACC 720

Qy     720 CCAGTGAAGGGG 731
Db     721 CCAGTGAAGGGG 732


RESULT 4
US-09-729-454-5/c
; Sequence 5, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 5
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137038A1 71851705V1
US-09-729-454-5
```

```
  Query Match            34.5%;  Score 558;  DB 10;  Length 621;
  Best Local Similarity  99.8%;  Pred. No. 1.7e-231;
  Matches  608;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    210 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 562

Qy    270 GGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGTCCTTAT 329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 GGTCTGCTCTAAACTTATCCATCTCCCTGCATTACTGGAACAACTCCACAAAGTCCTTAT 502

Qy    330 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 442

Qy    390 AAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGAACACCA 449
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    441 AAATCATAGAGAAACTAGATCAGCAGACCCCACCCAGACCTTTCACCCACGTGAACACCA 382

Qy    450 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 322

Qy    510 GAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 GAGAGCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 262

Qy    570 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 202

Qy    630 AGGTGACTGACTTGAACAACGGCACGTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 AGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 142

Qy    690 AGGTCTCTCTGTCTCTGCTGGTCATGCACCCGAGTGAAGGGGTGTCAGCTCTCTGGAGTG 749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 AGGTGTCTCTGTCTCTGCTGGTCATGCACCCGAGTGAAGGGGTGTCAGCTCTCTGGAGTG 82

Qy    750 CAAGGAACCAAGGCTATGACAGGGTGATCTTGAGTGGCCAGTTTGTCAATGGCACTTCCC 809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 CAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCC 22

Qy    810 AAGTCCACT 818
          |||||||||
Db     21 AAGTCCACT 13
```

```
RESULT 5
US-09-981-353-145
; Sequence 145, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 145
;   LENGTH: 841
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 403560.1
US-09-981-353-145

  Query Match            30.2%;  Score 488;  DB 9;  Length 841;
  Best Local Similarity  99.8%;  Pred. No. 2.6e-201;
  Matches  538;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1057 CAACACACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTTTGT 1116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CAACACACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTTTGT 362

Qy   1117 GGATTTGGATAGGAACATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATC 1176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GGATTTGGATAGGAACATCAACATCCAGTGGCAAAAACATTGTTATCCCTTGATAGGATC 422

Qy   1177 AATGACCTATTCAGTCAAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGG 1236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 AATGACCTATTCAGTCAAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGG 482

Qy   1237 AGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGA 1296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 AGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGA 542

Qy   1297 TGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAG 1356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAG 602

Qy   1357 CCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGA 1416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 CCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGA 662

Qy   1417 AAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGA 1476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    663 AAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGA 722

Qy   1477 TCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGT 1536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 TCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGT 782

Qy   1537 ACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTT 1595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    783 ACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTT 841
```

```
RESULT 6
US-09-729-454-6/c
; Sequence 6, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 6
;   LENGTH: 545
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137038A1 70255975V1
US-09-729-454-6

  Query Match            25.3%;  Score 409;  DB 10;  Length 545;
  Best Local Similarity  99.6%;  Pred. No. 3.3e-167;
  Matches  509;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    319 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 452

Qy    379 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCGATCACCCA 392
```

```
Qy     439 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 498
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     391 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 332

Qy     499 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 558
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     331 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 272

Qy     559 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 618
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 212

Qy     619 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 678
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 152

Qy     679 CTGGCAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 738
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 92

Qy     739 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 798
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 32

Qy     799 TGGCACTTCCCAAGTCCACTCTGAATGTGGC 829
           ||||||||||||||||||||||||||||||
Db      31 TGGCACTTCCCAAGTCCACTCTGAATGTGGC 1


RESULT 7
US-09-729-454-9
; Sequence 9, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 9
;    LENGTH: 394
     TYPE: DNA
     ORGANISM: Homo sapiens
     FEATURE:
     NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020137038A1 7128544H1
US-09-729-454-9

  Query Match           23.2%;  Score 375;  DB 10;  Length 394;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-152;
  Matches  375;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1221 TTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCA 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20 TTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCA 79

Qy    1281 GACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGC 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 GACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGC 139

Qy    1341 ATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGA 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 ATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGA 199

Qy    1401 TGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAA 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 TGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAA 259

Qy    1461 AGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCAT 1520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 AGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCAT 319

Qy    1521 ATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTAT 1580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 ATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTAT 379

Qy    1581 TAAACTATATTTGTT 1595
           |||||||||||||||
Db     380 TAAACTATATTTGTT 394


RESULT 8
US-09-729-454-4
; Sequence 4, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 4
;    LENGTH: 240
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020137038A1 2771041H1
US-09-729-454-4

  Query Match           14.9%;  Score 240;  DB 10;  Length 240;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-94;
  Matches  240;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240


RESULT 9
US-09-729-454-7
; Sequence 7, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 7
;    LENGTH: 236
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137038A1 5596934H1
;   NAME/KEY: unsure
;   LOCATION: 228
;   OTHER INFORMATION: a, t, c, g, or other
US-09-729-454-7

  Query Match           13.2%;  Score 213;  DB 10;  Length 236;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-82;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    766 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 74

Qy    826 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 134

Qy    886 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 194

Qy    946 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 978
          |||||||||||||||||||||||||||||||||
Db    195 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 227


RESULT 10
US-10-046-935-181
; Sequence 181, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 181
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-046-935-181

  Query Match           12.3%;  Score 198;  DB 9;  Length 554;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-76;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    864 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 60

Qy    924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 120

Qy    984 AGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTA 1043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTA 180

Qy   1044 GTGTCTCCAAATGCAACA 1061
          ||||||||||||||||||
Db    181 GTGTCTCCAAATGCAACA 198


RESULT 11
US-10-046-935-388/c
; Sequence 388, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 388
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 546, 560
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-388

  Query Match           5.8%;  Score 93;  DB 9;  Length 580;
  Best Local Similarity  99.3%;  Pred. No. 7.3e-31;
  Matches  143;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1062 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 85

Qy   1122 TGGATAGGAACATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGA 1181
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     84 TGGATAGGAACATCAACATCCAGTGGCAAAAACATTGTTATCCCTTGATAGGATCAATGA 25

Qy   1182 CCTATTCAGTCAAAGAGATGGAGT 1205
          |||||||||||||||||||||||||
Db     24 CCTATTCAGTCAAAGAGATGGAGT 1


RESULT 12
US-09-729-454-11
; Sequence 11, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 11
;   LENGTH: 661
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137038A1 1333949F6
US-09-729-454-11

  Query Match           5.4%;  Score 88;  DB 10;  Length 661;
  Best Local Similarity  100.0%;  Pred. No. 1e-28;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 403

Qy     411 AGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||
Db     404 AGCAGATCCCACCCAGACCTTTCACCCA 431


RESULT 13
US-09-729-454-31
; Sequence 31, Application US/09729454
  Patent No. US20020137038A1
  GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
   CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 31
;   LENGTH: 1545
    TYPE: DNA
;   ORGANISM: Homo sapiens
    FEATURE:
;   NAME/KEY: misc_feature
    OTHER INFORMATION: Incyte ID No. US20020137038A1 GNN.g9795680_006.edit
US-09-729-454-31

  Query Match             5.4%;  Score 88;  DB 10;  Length 1545;
  Best Local Similarity   100.0%;  Pred. No. 1e-28;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 145

Qy     411 AGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||
Db     146 AGCAGATCCCACCCAGACCTTTCACCCA 173


RESULT 14
US-09-729-454-10
; Sequence 10, Application US/09729454
; Patent No. US20020137038A1
  GENERAL INFORMATION:
   APPLICANT:  Lasek, Amy W.
   APPLICANT:  Krasnow, Randi E.
   APPLICANT:  Baughn, Mariah R.
   TITLE OF INVENTION:  INTESTINAL PROTEINS
   FILE REFERENCE:  PC-0028 CIP
   CURRENT APPLICATION NUMBER: US/09/729,454
   CURRENT FILING DATE:  2000-12-04
   NUMBER OF SEQ ID NOS:  32
   SOFTWARE:  PERL Program
  SEQ ID NO 10
    LENGTH: 2248
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    OTHER INFORMATION: Incyte ID No. US20020137038A1 7484349CB1
S-09-729-454-10

 Query Match             5.4%;  Score 88;  DB 10;  Length 2248;
 Best Local Similarity   100.0%;  Pred. No. 1e-28;
 Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Y      351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 403

Qy     411 AGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||
Db     404 AGCAGATCCCACCCAGACCTTTCACCCA 431


RESULT 15
US-09-729-454-8
; Sequence 8, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Krasnow, Randi E.
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION:  INTESTINAL PROTEINS
;  FILE REFERENCE:  PC-0028 CIP
;  CURRENT APPLICATION NUMBER: US/09/729,454
;  CURRENT FILING DATE:  2000-12-04
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PERL Program
; SEQ ID NO 8
;   LENGTH: 414
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137038A1 3229449F6
;   NAME/KEY: unsure
;   LOCATION: 47, 105, 248
;   OTHER INFORMATION: a, t, c, g, or other
US-09-729-454-8

  Query Match             4.5%;  Score 73;  DB 10;  Length 414;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-22;
  Matches   73;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1057 CAACACACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGT 1116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 CAACACACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGT 234

Qy    1117 GGATTTGGATAGG 1129
           |||||||||||||
Db     235 GGATTTGGATAGG 247


Search completed: November 22, 2002, 04:45:20
Job time : 873 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 01:11:26 ; Search time 4129 Seconds
(without alignments)
11390.200 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence: 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
42: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1616	100.0	1616	6	AX458378	AX458378 Sequence
2	1005	62.2	1005	6	AX458405	AX458405 Sequence
3	987	61.1	1739	6	AX358758	AX358758 Sequence
4	987	61.1	1739	6	AX362251	AX362251 Sequence
5	983	60.8	1138	6	AX083510	AX083510 Sequence
c 6	738	45.7	81975	2	AP000450	AP000450 Homo sapi
c 7	738	45.7	146841	9	AC020549	AC020549 Homo sapi
8	738	45.7	187852	2	AC068041	AC068041 Homo sapi
c 9	558	34.5	621	6	AX458380	AX458380 Sequence
10	528	32.7	1704	9	AK000134	AK000134 Homo sapi
c 11	409	25.3	545	6	AX458381	AX458381 Sequence
12	400	24.8	400	6	AX386470	AX386470 Sequence
13	375	23.2	394	6	AX458384	AX458384 Sequence
c 14	350	21.7	178494	2	AC067812	AC067812 Homo sapi
15	240	14.9	240	6	AX458379	AX458379 Sequence
16	213	13.2	236	6	AX458382	AX458382 Sequence
17	198	12.3	557	6	AX339934	AX339934 Sequence
18	192	11.9	275	6	AX079818	AX079818 Sequence
c 19	181	11.2	532	6	AX351427	AX351427 Sequence
c 20	93	5.8	583	6	AX340141	AX340141 Sequence
21	88	5.4	211	6	AX385391	AX385391 Sequence
22	88	5.4	611	6	AX385432	AX385432 Sequence
23	88	5.4	661	6	AX458386	AX458386 Sequence
24	88	5.4	1545	6	AX458406	AX458406 Sequence
25	88	5.4	1953	9	BC029049	BC029049 Homo sapi
26	88	5.4	2248	6	AX458385	AX458385 Sequence
27	88	5.4	81975	2	AP000450	AP000450 Homo sapi
28	83	5.1	2222	9	AK057953	AK057953 Homo sapi
29	83	5.1	143590	9	AP002774	AP002774 Homo sapi
30	73	4.5	414	6	AX458383	AX458383 Sequence
c 31	47	2.9	518	6	AX458387	AX458387 Sequence
32	42	2.6	2144	4	OCUNKNPA	Z12840 O.cuniculus
33	34	2.1	244	6	AX458393	AX458393 Sequence
34	34	2.1	270	6	AX458394	AX458394 Sequence
c 35	34	2.1	291	6	AX458392	AX458392 Sequence
c 36	33	2.0	60942	2	AC124611	AC124611 Mus muscu
37	32	2.0	206	6	AX458391	AX458391 Sequence
38	32	2.0	257	6	AX458397	AX458397 Sequence
39	29	1.8	58145	2	AC125828	AC125828 Rattus no
40	29	1.8	178494	2	AC067812	AC067812 Homo sapi
41	25	1.5	288	6	AX458395	AX458395 Sequence
42	24	1.5	275	6	AX458396	AX458396 Sequence
c 43	24	1.5	394	6	AX458384	AX458384 Sequence
c 44	24	1.5	150489	2	AC068854	AC068854 Homo sapi
45	24	1.5	165028	9	AC110758	AC110758 Homo sapi

ALIGNMENTS

RESULT 1
AX458378
LOCUS AX458378 1616 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 3 from Patent WO0246411.
ACCESSION AX458378
VERSION AX458378.1 GI:21725046
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lasek,A.W., Krasnow,R. and Baughn,M.R.
TITLE Intestinal proteins
JOURNAL Patent: WO 0246411-A 3 13-JUN-2002;

```
           Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1616
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3229449CB1"
BASE COUNT      511 a    374 c    324 g    407 t
ORIGIN

  Query Match             100.0%;  Score 1616;  DB 6;  Length 1616;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1616;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240

Qy     241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300

Qy     301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360

Qy     361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420

Qy     421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480

Qy     481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAG 540

Qy     541 GGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTC 600

Qy     601 CCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCT 660

Qy     661 GGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCC 720

Qy     721 CAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTT 780

Qy     781 CACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAA 840

Qy     841 CACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGT 900
```

```
Qy     901 GAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGT 960

Qy     961 TTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGAT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGAT 1020

Qy    1021 TATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTGGATCT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTGGATCT 1080

Qy    1081 GCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATCAACAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATCAACAT 1140

Qy    1141 CCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAAGAGAT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAAGAGAT 1200

Qy    1201 GGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTAT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTAT 1260

Qy    1261 TTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAA 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAA 1320

Qy    1321 TGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAA 1380

Qy    1381 AACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTA 1440

Qy    1441 CATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGC 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGC 1500

Qy    1501 CTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGG 1560

Qy    1561 AAATCAGATTAATATATTATTAAACTATATTTGTTAAATAACAAAAAAAAAAAAAA 1616
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AAATCAGATTAATATATTATTAAACTATATTTGTTAAATAACAAAAAAAAAAAAAA 1616
```

```
RESULT 2
AX458405
LOCUS       AX458405               1005 bp    DNA     linear   PAT 08-JUL-2002
DEFINITION  Sequence 30 from Patent WO0246411.
ACCESSION   AX458405
VERSION     AX458405.1  GI:21725073
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 30 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1005
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: GNN.g9965027_000007_006"
BASE COUNT      303 a    225 c    211 g    266 t
ORIGIN
```

```
  Query Match             62.2%;  Score 1005;  DB 6;  Length 1005;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1005;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     593 ATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGC 652
Db       1 ATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGC 60

Qy     653 ACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTC 712
Db      61 ACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTC 120

Qy     713 ATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGG 772
Db     121 ATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGG 180

Qy     773 GTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTG 832
Db     181 GTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTG 240

Qy     833 ATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTC 892
Db     241 ATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTC 300

Qy     893 TACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAAC 952
Db     301 TACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAAC 360

Qy     953 AAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGT 1012
Db     361 AAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGT 420

Qy    1013 GTAGAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCA 1072
Db     421 GTAGAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCA 480

Qy    1073 GTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAAC 1132
Db     481 GTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAAC 540

Qy    1133 ATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTC 1192
Db     541 ATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTC 600

Qy    1193 AAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTC 1252
Db     601 AAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTC 660

Qy    1253 ATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGG 1312
Db     661 ATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGG 720

Qy    1313 GCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTT 1372
Db     721 GCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTT 780

Qy    1373 ATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTT 1432
Db     781 ATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTT 840

Qy    1433 CATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATC 1492
Db     841 CATGGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTGTGAGTATC 900

Qy    1493 ATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACAT 1552
Db     901 ATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACACCCACCTCAACAT 960

Qy    1553 GTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTTAA 1597
Db     961 GTAGTCGGAAATCAGATTAATATATTATTAAACTATATTTGTTAA 1005
```

```
RESULT 3
AX358758
LOCUS       AX358758               1739 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 11 from Patent WO0193983.
ACCESSION   AX358758
VERSION     AX358758.1  GI:18675280
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baker,K.P., Desnoyers,L., Gerritsen,M.E., Goddard,A.,
            Godowski,P.J., Grimaldi,J.C., Gurney,A.L., Smith,V., Stephan,J.P.,
            Watanabe,C.K. and Wood,W.I.
  TITLE     Secreted and transmembrane polypeptides and nucleic acids encoding
            the same
  JOURNAL   Patent: WO 0193983-A 11 13-DEC-2001;
            Genentech Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1739
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      545 a    405 c    359 g    430 t
ORIGIN

  Query Match             61.1%;  Score 987;  DB 6;  Length 1739;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  987;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 134
Db       3 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 62

Qy     135 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 194
Db      63 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 122

Qy     195 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 254
Db     123 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 182

Qy     255 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 314
Db     183 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 242

Qy     315 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 374
Db     243 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 302

Qy     375 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 434
Db     303 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 362

Qy     435 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 494
Db     363 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 422

Qy     495 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 554
Db     423 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 482

Qy     555 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 614
Db     483 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 542

Qy     615 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 674
Db     543 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 602

Qy     675 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 734
Db     603 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 662
```

```
Qy     735 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 722

Qy     795 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 782

Qy     855 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 842

Qy     915 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 902

Qy     975 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 1034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 962

Qy    1035 ATACAATTAGTGTCTCCAAATGCAACA 1061
           ||||||||||||||||||||||||||
Db     963 ATACAATTAGTGTCTCCAAATGCAACA 989
```

```
RESULT 4
AX362251
LOCUS       AX362251               1739 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 11 from Patent WO0208288.
ACCESSION   AX362251
VERSION     AX362251.1  GI:18694583
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baker,K.P., Desnoyers,L., Gerritsen,M.E., Goddard,A.,
            Godowski,P.J., Grimaldi,J.C., Gurney,A.L., Smith,V., Stephan,J.P.,
            Watanabe,C.K. and Wood,W.I.
  TITLE     Secreted and transmembrane polypeptides and nucleic acids encoding
            the same
  JOURNAL   Patent: WO 0208288-A 11 31-JAN-2002;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1739
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      545 a    405 c    359 g    430 t
ORIGIN
```

```
  Query Match           61.1%;  Score 987;  DB 6;  Length 1739;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  987;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 62

Qy     135 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 122

Qy     195 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 182

Qy     255 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 242

Qy     315 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 302

Qy     375 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 362

Qy     435 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 422

Qy     495 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 482

Qy     555 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 542

Qy     615 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 602

Qy     675 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 662

Qy     735 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 722

Qy     795 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 782

Qy     855 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 842

Qy     915 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 902

Qy     975 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 1034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 962

Qy    1035 ATACAATTAGTGTCTCCAAATGCAACA 1061
           ||||||||||||||||||||||||||
Db     963 ATACAATTAGTGTCTCCAAATGCAACA 989
```

```
RESULT 5
AX083510
LOCUS       AX083510               1138 bp    DNA     linear   PAT 28-FEB-2001
DEFINITION  Sequence 52 from Patent WO0112662.
ACCESSION   AX083510
VERSION     AX083510.1  GI:13185320
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1138)
  AUTHORS   Lal,P., Yue,H., Tang,Y.T., Bandman,O., Burford,N., Azimzai,Y.,
            Baughn,M.R., Lu,D.A. and Patterson,C.
  TITLE     Membrane associated proteins
  JOURNAL   Patent: WO 0112662-A 52 22-FEB-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1138
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /note="Incyte ID No: 1334051CB1"
BASE COUNT       334 a    274 c    241 g    289 t
ORIGIN

  Query Match           60.8%;  Score 983;  DB 6;  Length 1138;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  983;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      19 CAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGTATGCCAGTAGCAC 78
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGTATGCCAGTAGCAC 60

Qy      79 ACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGAT 138
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGAT 120

Qy     139 TGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTC 198
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTC 180

Qy     199 ACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTC 258
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTC 240

Qy     259 CACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCAC 318
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCAC 300

Qy     319 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 378
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 360

Qy     379 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 420

Qy     439 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 498
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 480

Qy     499 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 558
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 540

Qy     559 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 618
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 600

Qy     619 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 678
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 660

Qy     679 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 738
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 720

Qy     739 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 798
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 780

Qy     799 TGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTG 858
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTG 840

Qy     859 CCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCC 918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCC 900

Qy     919 CTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 960

Qy     979 AGAAAAGAGCCTCTTTGAAAGGT 1001
           |||||||||||||||||||||||
Db     961 AGAAAAGAGCCTCTTTGAAAGGT 983


RESULT 6
AP000450/c
LOCUS       AP000450               81975 bp    DNA     linear   HTG 30-MAY-2000
DEFINITION  Homo sapiens chromosome 11 clone XXP1-307 map 11q23, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AP000450
VERSION     AP000450.3  GI:8118763
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens DNA, clone:XXP1-307.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 81975)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seoq,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 81,975 genomic DNA of 11q23
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 81975)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seoq,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     On May 31, 2000 this sequence version replaced gi:7630227.
            -------------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
                Center project name: HumDraft11
                Center clone name: XXP1-307
            -------------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 79702 bases at least Q40
                Consensus quality: 80980 bases at least Q30
                Consensus quality: 81428 bases at least Q20
                Insert size: 81575; sum-of-contigs
                Quality coverage: 5.78x in Q20 bases; sum-of-contigs
            --------------.
            NOTE: This is a 'working draft' sequence. It currently consists of
            5 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
                  1     35184 contig of  35184 bp in length
              35285     51779 contig of  16495 bp in length
              51880     63010 contig of  11131 bp in length
              63111     73886 contig of  10776 bp in length
              73987     81975 contig of   7989 bp in length
            Sequence updated (18-Apr-2000)
            Sequence updated (26-May-2000).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    35184: contig of 35184 bp in length
```

```
            *    35185 35284: gap of      100 bp
            *    35285    51779: contig of 16495 bp in length
            *    51780 51879: gap of      100 bp
            *    51880    63010: contig of 11131 bp in length
            *    63011 63110: gap of      100 bp
            *    63111    73886: contig of 10776 bp in length
            *    73887 73986: gap of      100 bp
            *    73987    81975: contig of 7989 bp in length.
FEATURES             Location/Qualifiers
     source          1. .81975
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q23"
                     /clone="XXP1-307"
     misc_feature    1. .35184
                     /note="assembly_fragment"
     misc_feature    35285. .51779
                     /note="assembly_fragment"
     misc_feature    51880. .63010
                     /note="assembly_fragment"
     misc_feature    63111. .73886
                     /note="assembly_fragment"
     misc_feature    73987. .81975
                     /note="assembly_fragment"
BASE COUNT    24973 a  15476 c  15771 g  25355 t    400 others
ORIGIN

  Query Match             45.7%;  Score 738;  DB 2;  Length 81975;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  738;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     264 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 323
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57709 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 57650

Qy     324 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 383
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57649 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 57590

Qy     384 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 443
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57589 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 57530

Qy     444 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 503
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57529 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 57470

Qy     504 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 563
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57469 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 57410

Qy     564 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 623
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57409 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 57350

Qy     624 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 683
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57349 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 57290

Qy     684 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 743
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57289 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 57230

Qy     744 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 803
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57229 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 57170

Qy     804 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 863
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57169 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 57110

Qy     864 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57109 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 57050

Qy     924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57049 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 56990

Qy     984 AGAGCCTCTTTGAAAGGT 1001
           ||||||||||||||||||
Db   56989 AGAGCCTCTTTGAAAGGT 56972


RESULT 7
AC020549/c
LOCUS       AC020549              146841 bp    DNA     linear   PRI 09-MAY-2001
DEFINITION  Homo sapiens BAC clone RP11-172C16 from 11, complete sequence.
ACCESSION   AC020549
VERSION     AC020549.4  GI:9795680
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 146841)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 146841)
  AUTHORS   Edwards,J., Drone,K., LaPlant,Y. and Ahluwalia,R.
  TITLE     The sequence of Homo sapiens BAC clone RP11-172C16
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 146841)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 146841)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 146841)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 146841)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Aug 11, 2000 this sequence version replaced gi:7630875
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0172C16
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-720G3; the clone sequenced to the right is RP11-643E23. Actual start of this clone is at base position 1 of RP11-172C16; actual end is at base position 146841 of RP11-172C16.

The region from base position 130951 to 131438 represents the best possible assembly of an imperfect TA repeat. Restriction digest information from EcoRV, HindIII, and BamHI indicates that approximately 150 to 250 bases may be missing from the final sequence.

FEATURES Location/Qualifiers
source 1. .146841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-172C16"
/clone_lib="RPCI-11"
repeat_region 418. .575
/rpt_family="MER1_type"
repeat_region 774. .946
/rpt_family="L1"
repeat_region 1114. .1313
/rpt_family="L1"
repeat_region 1365. .1452
/rpt_family="ERV"
repeat_region 1425. .1510
/rpt_family="ERV"
repeat_region 1541. .2262
/rpt_family="L1"
repeat_region 2269. .2957
/rpt_family="L1"
repeat_region 8426. .8655
/rpt_family="L1"
repeat_region 8656. .8957
/rpt_family="Alu"
repeat_region 8958. .8996
/rpt_family="L1"
repeat_region 9029. .9453
/rpt_family="L1"
repeat_region 10059. .10457
/rpt_family="MaLR"
repeat_region 10468. .10610
/rpt_family="L1"
repeat_region 10659. .10807
/rpt_family="L2"
repeat_region 10843. .11005
/rpt_family="ERVL"
repeat_region 12544. .12725
/rpt_family="L1"
repeat_region 12830. .13126
/rpt_family="Alu"
repeat_region 13433. .13674
/rpt_family="L1"
repeat_region 14419. .14587
/rpt_family="MIR"
repeat_region 15199. .15545
/rpt_family="L2"
repeat_region 16012. .16077
/rpt_family="L2"
misc_feature 17372. .17741
/note="similar to EST AW009077 (NID:g5857855) ws76q06.x1"
misc_feature 17555. .18180
/note="similar to EST AW844097 (NID:g7938080)"
misc_feature 17696. .18060
/note="similar to EST AW841802 (NID:g7935785)"
repeat_region 17999. .18157
/rpt_family="L2"
repeat_region 18220. .18324
/rpt_family="L1"
repeat_region 18325. .18401
/rpt_family="Alu"
repeat_region 18408. .18538
/rpt_family="Alu"
repeat_region 18539. .19181
/rpt_family="L1"
repeat_region 19454. .19661
/rpt_family="L1"
repeat_region 19705. .20140
/rpt_family="L1"
misc_feature 20162. .20624
/note="similar to EST AI347274 (NID:g4084480) tc05d02.x1"
misc_feature 20371. .20690
/note="similar to EST AW844024 (NID:g7938007)"
repeat_region 22353. .22774
/rpt_family="MaLR"
repeat_region 23334. .23715
/rpt_family="ERVL"
repeat_region 23738. .23802
/rpt_family="MaLR"
repeat_region 23803. .24380
/rpt_family="ERV1"
repeat_region 24381. .24774
/rpt_family="MaLR"
repeat_region 24925. .25100
/rpt_family="L2"
repeat_region 25169. .25250
/rpt_family="MER53"
repeat_region 25524. .25566
/rpt_family="L2"
repeat_region 25567. .25668
/rpt_family="U13"
repeat_region 26564. .26944
/rpt_family="L2"
repeat_region 26947. .27061
/rpt_family="L2"
repeat_region 28123. .28317
/rpt_family="L2"
repeat_region 29799. .30188
/rpt_family="L2"
repeat_region 30312. .30484
/rpt_family="MER1_type"
repeat_region 38493. .38692
/rpt_family="Alu"
repeat_region 38712. .38863
/rpt_family="L1"
repeat_region 39064. .39272
/rpt_family="L1"
repeat_region 39323. .39868

Query Match 45.7%; Score 738; DB 9; Length 146841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     264 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 323
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81461 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 81402

Qy     324 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 383
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81401 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 81342

Qy     384 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 443
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81341 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 81282

Qy     444 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 503
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81281 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 81222

Qy     504 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 563
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81221 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 81162

Qy     564 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 623
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81161 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 81102

Qy     624 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 683
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81101 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 81042

Qy     684 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 743
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81041 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 80982

Qy     744 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 803
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80981 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 80922

Qy     804 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 863
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80921 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 80862

Qy     864 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80861 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 80802

Qy     924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80801 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 80742

Qy     984 AGAGCCTCTTTGAAAGGT 1001
           ||||||||||||||||||
Db   80741 AGAGCCTCTTTGAAAGGT 80724


RESULT 8
AC068041
LOCUS       AC068041              187852 bp    DNA     linear   HTG 02-SEP-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-643E23, WORKING DRAFT
            SEQUENCE, 9 unordered pieces.
ACCESSION   AC068041
VERSION     AC068041.4  GI:9965027
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187852)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 187852)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Sep 2, 2000 this sequence version replaced gi:8469045.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0643E23
            -------------- Summary Statistics --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 183822 bases at least Q40
            Consensus quality: 185080 bases at least Q30
            Consensus quality: 185802 bases at least Q20
            Insert size: 186000; agarose-fp
            Insert size: 187052; sum-of-contigs
            Quality coverage: 7.61 in Q20 bases; agarose-fp
            Quality coverage: 7.02 in Q20 bases; sum-of-contigs
            ------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1761: contig of 1761 bp in length
            *     1762     1861: gap of unknown length
            *     1862     8334: contig of 6473 bp in length
            *     8335     8434: gap of unknown length
            *     8435    15913: contig of 7479 bp in length
            *    15914    16013: gap of unknown length
            *    16014    23102: contig of 7089 bp in length
            *    23103    23202: gap of unknown length
            *    23203    31155: contig of 7953 bp in length
            *    31156    31255: gap of unknown length
            *    31256    45995: contig of 14740 bp in length
            *    45996    46095: gap of unknown length
            *    46096    77306: contig of 31211 bp in length
            *    77307    77406: gap of unknown length
            *    77407   144505: contig of 67099 bp in length
            *   144506   144605: gap of unknown length
            *   144606   187852: contig of 43247 bp in length.
FEATURES             Location/Qualifiers
     source          1. .187852
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP11-643E23"
     misc_feature    1. .1761
                     /note="assembly_name:Contig34"
     misc_feature    1862. .8334
                     /note="assembly_name:Contig35"
     misc_feature    8435. .15913
                     /note="assembly_name:Contig36"
     misc_feature    16014. .23102
                     /note="assembly_name:Contig37"
     misc_feature    23203. .31155
                     /note="assembly_name:Contig38"
     misc_feature    31256. .45995
                     /note="assembly_name:Contig39"
     misc_feature    46096. .77306
                     /note="assembly_name:Contig40
                     clone_end:SP6
                     vector_side:right"
     misc_feature    77407. .144505
                     /note="assembly_name:Contig42
```

```
                     clone_end:T7
                     vector_side:right"
     misc_feature    144606. .187852
                     /note="assembly_name:Contig41"
BASE COUNT     55964 a  35742 c  36408 g  58936 t    802 others
ORIGIN

  Query Match           45.7%;  Score 738;  DB 2;  Length 187852;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  738;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     264 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 323
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64776 AGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGT 64835

Qy     324 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 383
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64836 CCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAA 64895

Qy     384 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 443
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64896 TAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGA 64955

Qy     444 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 503
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64956 ACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACT 65015

Qy     504 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 563
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65016 GCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGC 65075

Qy     564 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 623
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65076 AATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTT 65135

Qy     624 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 683
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65136 CAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGG 65195

Qy     684 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 743
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65196 AGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCT 65255

Qy     744 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 803
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65256 GGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCA 65315

Qy     804 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 863
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65316 CTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGT 65375

Qy     864 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65376 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 65435

Qy     924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65436 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 65495

Qy     984 AGAGCCTCTTTGAAAGGT 1001
           ||||||||||||||||||
Db   65496 AGAGCCTCTTTGAAAGGT 65513


ESULT 9
X458380/c
OCUS       AX458380               621 bp    DNA     linear   PAT 08-JUL-2002
EFINITION  Sequence 5 from Patent WO0246411.
CCESSION   AX458380
ERSION     AX458380.1  GI:21725048
EYWORDS    .
OURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 5 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 71851705V1"
BASE COUNT      142 a    138 c    179 g    162 t
ORIGIN

  Query Match           34.5%;  Score 558;  DB 6;  Length 621;
  Best Local Similarity  99.8%;  Pred. No. 1.4e-305;
  Matches  608;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     210 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 269
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     621 TGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCACAAAGGTTT 562

Qy     270 GGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGTCCTTAT 329
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     561 GGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCACAAAGTCCTTAT 502

Qy     330 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 389
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     501 TCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGG 442

Qy     390 AAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCACGTGAACACCA 449
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     441 AAATCATAGAGAAACTAGATCAGCAGACCCCACCCAGACCTTTCACCCACGTGAACACCA 382

Qy     450 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 509
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     381 CCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACGTACTGCAGGG 322

Qy     510 GAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 569
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 GAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAGGAAGCAATATG 262

Qy     570 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 GCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAA 202

Qy     630 AGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 689
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     201 AGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCC 142

Qy     690 AGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTG 749
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     141 AGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTG 82

Qy     750 CAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCC 809
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 CAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCC 22

Qy     810 AAGTCCACT 818
           |||||||||
Db      21 AAGTCCACT 13


RESULT 10
AK000134
LOCUS       AK000134               1704 bp    mRNA    linear   PRI 22-FEB-2000
DEFINITION  Homo sapiens cDNA FLJ20127 fis, clone COL06176.
ACCESSION   AK000134
VERSION     AK000134.1  GI:7020022
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL06176.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1704)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1,
            Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .1704
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL06176"
                     /tissue_type="colon"
                     /clone_lib="COL"
                     /note="cloning vector pME18SFL3"
     CDS             558. .1040
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA90965.1"
                     /db_xref="GI:7020023"
                     /translation="MEKFNTISASKCNKETVAMKEKCKFGMTSTIPSGHVWRNTWNPV
                     SCSLATVKMKECLRGKLIYLMGDSTIRQWMEYFKASINTLKSVDLHESGKLQHQLAVD
                     LDRNINIQWQKYCYPLIGSMTYSVKEMEYLTRAIDRTGGEKKYCHCYFPGPAFQTLSH
                     "
BASE COUNT      597 a    366 c    308 g    433 t
ORIGIN

  Query Match           32.7%;  Score 528;  DB 9;  Length 1704;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-288;
  Matches  528;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240

Qy     241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCA 300

Qy     301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCC 360

Qy     361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCC 420

Qy     421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCAT 480

Qy     481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CCTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCT 528


RESULT 11
AX458381/c
LOCUS       AX458381                545 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 6 from Patent WO0246411.
ACCESSION   AX458381
VERSION     AX458381.1  GI:21725049
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 6 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 70255975V1"
BASE COUNT      110 a    129 c    161 g    145 t
ORIGIN

  Query Match           25.3%;  Score 409;  DB 6;  Length 545;
  Best Local Similarity  99.6%;  Pred. No. 9.9e-221;
  Matches  509;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     319 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     511 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 452

Qy     379 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCGATCACCCA 392

Qy     439 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     391 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 332

Qy     499 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     331 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 272

Qy     559 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 212

Qy     619 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 152

Qy     679 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 92

Qy     739 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 32

Qy     799 TGGCACTTCCCAAGTCCACTCTGAATGTGGC 829
          ||||||||||||||||||||||||||||||
```

```
Db      31 TGGCACTTCCCAAGTCCACTCTGAATGTGGC 1


RESULT 12
AX386470
LOCUS       AX386470                400 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1398 from Patent WO0214500.
ACCESSION   AX386470
VERSION     AX386470.1  GI:19579600
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
            Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
            Labat,I.
  TITLE     Human genes and gene expression products
  JOURNAL   Patent: WO 0214500-A 1398 21-FEB-2002;
            CHIRON CORPORATION (US) ; Hyseq Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      110 a     98 c     92 g    100 t
ORIGIN

  Query Match             24.8%;  Score 400;  DB 6;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-215;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     646 CAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCT 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCT 60

Qy     706 GCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTA 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTA 120

Qy     766 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 180

Qy     826 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 240

Qy     886 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 300

Qy     946 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAA 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAA 360

Qy    1006 TGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTAGT 1045
           ||||||||||||||||||||||||||||||||||||||||
Db     361 TGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTAGT 400


RESULT 13
AX458384
LOCUS       AX458384                394 bp    DNA     linear   PAT 08-JUL-2002
DEFINITION  Sequence 9 from Patent WO0246411.
ACCESSION   AX458384
VERSION     AX458384.1  GI:21725052
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 9 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7128544H1"
BASE COUNT      126 a     80 c     71 g    117 t
ORIGIN

  Query Match             23.2%;  Score 375;  DB 6;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-201;
  Matches  375;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1221 TTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCA 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20 TTGACAGAACTGGAGGAGAAAAAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCA 79

Qy    1281 GACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGC 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 GACCCTTTCCCATTGATGTTTTTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGC 139

Qy    1341 ATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGA 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 ATCTTCTTCTGAGAAGCCCAGACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGA 199

Qy    1401 TGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAA 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 TGTACAATGATGCAGAAAGATTTAGTGACTTTCATGGTTACATTCAATATCTCATCATAA 259

Qy    1461 AGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCAT 1520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 AGGACATTTTCCAGGATCTCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCAT 319

Qy    1521 ATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTAT 1580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 ATGGCACAAATAATGTACACCCACCTCAACATGTAGTCGGAAATCAGATTAATATATTAT 379

Qy    1581 TAAACTATATTTGTT 1595
           |||||||||||||||
Db     380 TAAACTATATTTGTT 394


RESULT 14
AC067812/c
LOCUS       AC067812              178494 bp    DNA     linear   HTG 04-MAY-2001
DEFINITION  Homo sapiens chromosome 11 clone RP11-720G3 map 11, WORKING DRAFT
            SEQUENCE, 32 unordered pieces.
ACCESSION   AC067812
VERSION     AC067812.2  GI:8318580
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 178494)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-720G3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 178494)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
```

```
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     on Jun 7, 2000 this sequence version replaced gi:7651863.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9810
                Center clone name: 720_G_3
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 161804 bases at least Q40
                Consensus quality: 170435 bases at least Q30
                Consensus quality: 173827 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 175394; sum-of-contigs
                Quality coverage: 3.6 in Q20 bases; agarose-fp
                Quality coverage: 3.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 32 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1247: contig of 1247 bp in length
            *     1248 1347: gap of      100 bp
            *     1348     2922: contig of 1575 bp in length
            *     2923 3022: gap of      100 bp
            *     3023     4326: contig of 1304 bp in length
            *     4327 4426: gap of      100 bp
            *     4427     5770: contig of 1344 bp in length
            *     5771 5870: gap of      100 bp
            *     5871     7537: contig of 1667 bp in length
            *     7538 7637: gap of      100 bp
            *     7638    10278: contig of 2641 bp in length
            *    10279 10378: gap of      100 bp
            *    10379    12229: contig of 1851 bp in length
            *    12230 12329: gap of      100 bp
            *    12330    14202: contig of 1873 bp in length
            *    14203 14302: gap of      100 bp
            *    14303    17175: contig of 2873 bp in length
            *    17176 17275: gap of      100 bp
            *    17276    20747: contig of 3472 bp in length
            *    20748 20847: gap of      100 bp
            *    20848    24235: contig of 3388 bp in length
            *    24236 24335: gap of      100 bp
            *    24336    28438: contig of 4103 bp in length
            *    28439 28538: gap of      100 bp
            *    28539    32248: contig of 3710 bp in length
            *    32249 32348: gap of      100 bp
            *    32349    35121: contig of 2773 bp in length
            *    35122 35221: gap of      100 bp
            *    35222    38838: contig of 3617 bp in length
            *    38839 38938: gap of      100 bp
            *    38939    43464: contig of 4526 bp in length
            *    43465 43564: gap of      100 bp
            *    43565    48133: contig of 4569 bp in length
            *    48134 48233: gap of      100 bp
            *    48234    52993: contig of 4760 bp in length
            *    52994 53093: gap of      100 bp
            *    53094    57397: contig of 4304 bp in length
            *    57398 57497: gap of      100 bp
            *    57498    62668: contig of 5171 bp in length
            *    62669 62768: gap of      100 bp
            *    62769    67716: contig of 4948 bp in length
            *    67717 67816: gap of      100 bp
            *    67817    74417: contig of 6601 bp in length
            *    74418 74517: gap of      100 bp
            *    74518    82397: contig of 7880 bp in length
            *    82398 82497: gap of      100 bp
            *    82498    89948: contig of 7451 bp in length
            *    89949 90048: gap of      100 bp
            *    90049    98408: contig of 8360 bp in length
            *    98409 98508: gap of      100 bp
            *    98509   108229: contig of 9721 bp in length
            *   108230 108329: gap of      100 bp
            *   108330   118955: contig of 10626 bp in length
            *   118956 119055: gap of      100 bp
            *   119056   129091: contig of 10036 bp in length
            *   129092 129191: gap of      100 bp
            *   129192   140666: contig of 11475 bp in length
            *   140667 140766: gap of      100 bp
            *   140767   150098: contig of 9332 bp in length
            *   150099 150198: gap of      100 bp
            *   150199   161240: contig of 11042 bp in length
            *   161241 161340: gap of      100 bp
            *   161341   178494: contig of 17154 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178494
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-720G3"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1247
                     /note="assembly_fragment"
     misc_feature    1348. .2922
                     /note="assembly_fragment"
     misc_feature    3023. .4326
                     /note="assembly_fragment"
     misc_feature    4427. .5770
                     /note="assembly_fragment"
     misc_feature    5871. .7537
                     /note="assembly_fragment"
     misc_feature    7638. .10278
                     /note="assembly_fragment"
     misc_feature    10379. .12229
                     /note="assembly_fragment"
     misc_feature    12330. .14202
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    14303. .17175
                     /note="assembly_fragment"
     misc_feature    17276. .20747
                     /note="assembly_fragment"
     misc_feature    20848. .24235
                     /note="assembly_fragment"
     misc_feature    24336. .28438
                     /note="assembly_fragment"
     misc_feature    28539. .32248
                     /note="assembly_fragment"
     misc_feature    32349. .35121
                     /note="assembly_fragment"
```

```
     misc_feature    35222. .38838
                     /note="assembly_fragment"
     misc_feature    38939. .43464
                     /note="assembly_fragment"
     misc_feature    43565. .48133
                     /note="assembly_fragment"
     misc_feature    48234. .52993
                     /note="assembly_fragment"
     misc_feature    53094. .57397
                     /note="assembly_fragment"
     misc_feature    57498. .62668
                     /note="assembly_fragment"
     misc_feature    62769. .67716
                     /note="assembly_fragment"
     misc_feature    67817. .74417
                     /note="assembly_fragment"
     misc_feature    74518. .82397
                     /note="assembly_fragment"
     misc_feature    82498. .89948
                     /note="assembly_fragment"
     misc_feature    90049. .98408
                     /note="assembly_fragment"
     misc_feature    98509. .108229
                     /note="assembly_fragment"
     misc_feature    108330. .118955
                     /note="assembly_fragment"
     misc_feature    119056. .129091
                     /note="assembly_fragment"

  Query Match             21.7%;  Score 350;  DB 2;  Length 178494;
  Best Local Similarity   99.8%;  Pred. No. 4.3e-187;
  Matches  400;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1062 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 1121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  150599 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 150540

Qy    1122 TGGATAGGAACATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGA 1181
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  150539 TGGATAGGAACATCAACATCCAGTGGCAAAAACATTGTTATCCCTTGATAGGATCAATGA 150480

Qy    1182 CCTATTCAGTCAAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAA 1241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  150479 CCTATTCAGTCAAAGAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAA 150420

Qy    1242 AAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTT 1301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  150419 AAAATACTGTCATTGTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTT 150360

Qy    1302 TTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAG 1361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  150359 TTATCCGAAGGGCCCTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAG 150300

Qy    1362 ACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGAT 1421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  150299 ACACTATGGTTATCATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGAT 150240

Qy    1422 TTAGTGACTTTCATGGTTACATTCAATATCTCATCATAAAG 1462
           |||||||||||||||||||||||||||||||||||||||||
Db  150239 TTAGTGACTTTCATGGTTACATTCAATATCTCATCATAAAG 150199


RESULT 15
AX458379
LOCUS       AX458379                 240 bp    DNA     linear   PAT 08-JUL-2002
DEFINITION  Sequence 4 from Patent WO0246411.
ACCESSION   AX458379
VERSION     AX458379.1  GI:21725047
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Cranlata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1
  AUTHORS   Lasek,A.W., Krasnow,R. and Baughn,M.R.
  TITLE     Intestinal proteins
  JOURNAL   Patent: WO 0246411-A 4 13-JUN-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .240
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 2771041H1"
BASE COUNT       81 a     42 c     44 g     73 t
ORIGIN

  Query Match             14.9%;  Score 240;  DB 6;  Length 240;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-124;
  Matches  240;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCATGCTAAAGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAA 60

Qy      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGTATGCCAGTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAG 120

Qy     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGTCTATGTATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAG 180

Qy     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTT 240


Search completed: November 22, 2002, 03:58:28
Job time : 4503 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 01:08:46 ; Search time 332 Seconds
(without alignments)
10961.534 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	987	61.1	1739	24	ABK33541	cDNA encoding huma
	2	983	60.8	1138	22	AAF81755	Human membrane ass
	3	880	54.5	1629	24	ABQ86149	Novel human gene.
	4	694	42.9	1137	23	AAS76631	DNA encoding novel
	5	400	24.8	400	24	ABN61431	Human cancer relat
	6	386	23.9	437	23	AAS92089	DNA encoding novel
	7	198	12.3	557	24	ABL36592	Human colon tumour
	8	192	11.9	275	22	AAF93741	cDNA encoding SRT
	9	182	11.3	817	22	AAK88279	Human digestive sy
c	10	181	11.2	532	24	ABK27737	Human colon cancer
c	11	93	5.8	583	24	ABL36799	Human colon tumour
	12	88	5.4	211	24	ABN60352	Human cancer relat
	13	88	5.4	611	24	ABN60393	Human cancer relat
	14	88	5.4	684	24	ABK69141	DNA encoding human
	15	88	5.4	1177	24	ABK69140	DNA encoding human
	16	88	5.4	1662	24	ABQ86150	Novel human gene.
	17	88	5.4	1773	23	AAS92091	DNA encoding novel
	18	88	5.4	2089	24	ABK69106	DNA encoding human
	19	60	3.7	60	24	ABN48360	Human spliced tran
	20	53	3.3	71	21	AAC30578	Human secreted pro
	21	29	1.8	3676	23	AAS68944	DNA encoding novel
	22	24	1.5	580	22	AAH98108	Murine 7-transmemb
	23	23	1.4	600	22	AAH34126	Human colon cancer
c	24	23	1.4	1562	22	AAF32540	Mouse P/Q-type cal
c	25	23	1.4	3021	22	AAF32539	Mouse P/Q-type cal
c	26	23	1.4	6542	22	AAF32538	Mouse P/Q-type cal
c	27	23	1.4	9373	24	ABL70560	Chemically treated
c	28	23	1.4	9373	24	ABL33291	Human immune syste
c	29	23	1.4	9373	24	AAS61218	Human gene regulat
c	30	23	1.4	9373	24	ABK31317	Signal transductio
c	31	23	1.4	14307	24	ABL32728	Human immune syste
	32	22	1.4	243	21	AAA64605	Nucleotide sequenc
c	33	22	1.4	345	21	AAC25014	Human secreted pro
c	34	22	1.4	396	22	AAF94849	Human ovarian canc
c	35	22	1.4	396	24	ABT03116	Human ovarian carc
c	36	22	1.4	396	24	ABL48799	Ovarian carcinoma
	37	22	1.4	548	23	ABV48712	Human prostate exp
	38	22	1.4	556	22	AAH11526	Human cDNA clone (
c	39	22	1.4	607	21	AAA16283	Human colon cancer
	40	22	1.4	656	23	ABV18933	Human prostate exp
	41	22	1.4	980	23	AAS93510	DNA encoding novel
	42	22	1.4	2007	21	AAZ44859	Plant pathogen ind
c	43	22	1.4	2012	24	ABK34756	Human cDNA for nov
c	44	22	1.4	2497	22	AAH14473	Human cDNA sequenc
	45	22	1.4	4158	22	AAC85098	Atherosclerosis-as

ALIGNMENTS

RESULT 1
ABK33541
ID ABK33541 standard; cDNA; 1739 BP.
XX
AC ABK33541;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA encoding human PRO protein, Seq ID No 11.
XX
KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200208288-A2.
XX
PD 31-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US21066.
XX
PR 20-JUL-2000; 2000US-219556P.
PR 25-JUL-2000; 2000US-220585P.
PR 25-JUL-2000; 2000US-220605P.
PR 25-JUL-2000; 2000US-220607P.
PR 25-JUL-2000; 2000US-220624P.
PR 25-JUL-2000; 2000US-220638P.
PR 25-JUL-2000; 2000US-220664P.
PR 25-JUL-2000; 2000US-220666P.
PR 26-JUL-2000; 2000US-220893P.

PR 28-JUL-2000; 2000WO-US20710.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 15-SEP-2000; 2000US-000000P.
PR 10-NOV-2000; 2000WO-US30873.
PR 28-NOV-2000; 2000US-253646P.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000US-0747259.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001WO-US17092.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2002-172001/22.
DR P-PSDB; AAU83597.
XX
PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for treating a PRO related disorder and for diagnosing tumours
PT such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT tumour or liver tumour -
XX
PS Claim 2; Figure 11; 359pp; English.
XX
CC The invention relates to one hundred and twenty two nucleic acids
CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
CC encode human secreted proteins. The PRO nucleic acids, polypeptides,
CC agonists and antagonists are useful for treating a PRO related disorder.
CC The PRO polypeptides are useful for diagnosing tumours, especially lung
CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. The PRO polypeptides are useful for stimulating the
CC proliferation of, or gene expression, in pericyte cells, for stimulating
CC the proliferation or differentiation of chondrocyte cells, for
CC stimulating the release of tumour necrosis factor-alpha from human blood,
CC for stimulating or inhibiting the proliferation of normal human dermal
CC fibroblast cells. The PRO polypeptide may also be used as molecular
CC weight markers and for tissue typing. The PRO nucleic acids have
CC applications in molecular biology, including use as hybridisation probes,
CC and in chromosome and gene mapping. ABK33536-ABK33657 represent human
CC PRO protein coding sequences of the invention.
XX
SQ Sequence 1739 BP; 545 A; 405 C; 359 G; 430 T; 0 other;

Query Match 61.1%; Score 987; DB 24; Length 1739;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      75 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 GCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATG 62

Qy     135 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATA 122

Qy     195 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGA 182

Qy     255 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 ACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACT 242

Qy     315 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 CCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTG 302

Qy     375 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCA 362

Qy     435 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 CCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAG 422

Qy     495 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGAC 482

Qy     555 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGG 542

Qy     615 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 CAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTC 602

Qy     675 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 TGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGT 662

Qy     735 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 CAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTG 722

Qy     795 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAAT 782

Qy     855 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 TGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACA 842

Qy     915 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 TGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCA 902

Qy     975 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 1034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCA 962

Qy    1035 ATACAATTAGTGTCTCCAAATGCAACA 1061
           |||||||||||||||||||||||||||
Db     963 ATACAATTAGTGTCTCCAAATGCAACA 989
```

RESULT 2
AAF81755
ID AAF81755 standard; cDNA; 1138 BP.
XX
AC AAF81755;
XX
DT 12-JUN-2001 (first entry)
XX
DE Human membrane associated protein MEMAP-15 encoding cDNA.
XX
KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW epilepsy; diarrhoea; ss.
XX
OS Homo sapiens.
XX
PN WO200112662-A2.
XX
PD 22-FEB-2001.
XX
PF 14-AUG-2000; 2000WO-US22315.
XX
PR 17-AUG-1999; 99US-0149641.

Filed 14 Aug 2000

```
PR   09-NOV-1999;   99US-0164203.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lal P,  Yue H,  Tang YT,  Bandman O,  Burford N,  Azimzai Y;
PI   Baughn MR,  Lu DAM,  Patterson C;
XX
DR   WPI; 2001-168860/17.
DR   P-PSDB; AAB74709.
XX
PT   Isolated polypeptide with a human membrane associated protein sequence
PT   is useful for the diagnosis, prevention and treatment of cell
PT   proliferative, autoimmune/inflammatory, neurological and
PT   gastrointestinal disorders -
XX
PS   Claim 5; Page 159; 173pp; English.
XX
CC   AAF81741 to AAF81777 encode the human membrane associated proteins
CC   (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
CC   antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and
CC   antiarteriosclerotic activities, which can be used in gene therapy.
CC   MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
CC   associated with decreased expression of functional MEMAP and antagonists
CC   of MEMAP are used to treat a disease or condition associated with
CC   overexpression of functional MEMAP. These disorders include cell
CC   proliferative, autoimmune/inflammatory, neurological and gastrointestinal
CC   disorders. The MEMAP polynucleotides and proteins are also used for the
CC   diagnosis of these disorders. Specific examples of these disorders
CC   include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
CC   MEMAP proteins can be used to screen for compounds which specifically
CC   bind MEMAP including antibodies, oligonucleotides, proteins and small
CC   molecules. MEMAP polynucleotides can be used to prepare transgenic
CC   animals which can be studied to provide information concerning human
CC   disease. Anti-MEMAP antibodies are useful in immunoassays for the
CC   detection of MEMAP protein and can be used as antagonists to treat or
CC   prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
CC   can be delivered to target cells with genetic abnormalities with respect
CC   to the expression of MEMAP to treat or prevent a disorder associated
CC   with MEMAP.
XX
SQ   Sequence 1138 BP; 334 A; 274 C; 241 G; 289 T; 0 other;

  Query Match             60.8%;  Score 983;  DB 22;  Length 1138;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  983;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      19 CAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGTATGCCAGTAGCAC 78
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGTATGCCAGTAGCAC 60

Qy      79 ACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGAT 138
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGAT 120

Qy     139 TGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTC 198
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTC 180

Qy     199 ACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTC 258
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTC 240

Qy     259 CACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCAC 318
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACAACTCCAC 300

Qy     319 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 378
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACT 360

Qy     379 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTTCACCCA 420

Qy     439 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 498
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATAC 480

Qy     499 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 558
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACTTGGGACGCAG 540

Qy     559 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 618
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCCCCAGCGCTGATGGCAGG 600

Qy     619 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 678
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTT 660

Qy     679 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 738
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCCAGTGAAGGGGTGTCAGC 720

Qy     739 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 798
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTCACTGGCCAGTTTGTCAA 780

Qy     799 TGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTG 858
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTG 840

Qy     859 CCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCC 918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCC 900

Qy     919 CTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACA 960

Qy     979 AGAAAAGAGCCTCTTTGAAAGGT 1001
           |||||||||||||||||||||||
Db     961 AGAAAAGAGCCTCTTTGAAAGGT 983


RESULT 3
ABQ86149
ID   ABQ86149 standard; DNA; 1629 BP.
XX
AC   ABQ86149;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Novel human gene. SEQ ID 20.
XX
KW   Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;
KW   nootropic; neuroprotective; immunosuppressive; haemostatic;
KW   antiinflammatory; cardiant; antiulcer; virucide; antithyroid;
KW   cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW   wound healing disorders; atherosclerosis; Parkinson's disease;
KW   Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW   inflammation; neoplastic disease; nervous system disorder;
KW   cardiovascular disorders; pancreatitis; respiratory disorder;
KW   hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW   developmental abnormality; gastrointestinal ulceration; neuropathy;
KW   haematological disease; metabolic disease; sperm dysfunction;
KW   thyroid disorder; hypothyroidism; brain damage; colitis;
KW   cone photo- transduction deficiency; neurological disease; stroke;
KW   angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW   trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW   growth abnormality; precocious puberty; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250105-A1.
XX
```

PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US49232.
XX
PR 19-DEC-2000; 2000US-256710P.
PR 20-DEC-2000; 2000US-257048P.
PR 09-JAN-2001; 2001US-260482P.
PR 30-JAN-2001; 2001US-264922P.
PR 06-FEB-2001; 2001US-266797P.
PR 19-MAR-2001; 2001US-276988P.
PR 04-APR-2001; 2001US-281535P.
PR 08-MAY-2001; 2001US-289622P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
DR WPI; 2002-508784/54.
DR P-PSDB; ABP60984.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune
PT disorder -
XX
PS Claim 2(a); Page 235; 335pp; English.
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated.
CC The activity of polypeptides of the invention may be described as,
CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
CC of the invention.
XX
SQ Sequence 1629 BP; 497 A; 391 C; 329 G; 412 T; 0 other;

Query Match 54.5%; Score 880; DB 24; Length 1629;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 182 ATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTT 241
Db 1 ATGATAAATTATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTT 60

Qy 242 ACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCAT 301
Db 61 ACAGTTTTCCAGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCAT 120

Qy 302 TACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCA 361
Db 121 TACTGGAACAACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCA 180

Qy 362 CTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCA 421
Db 181 CTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCA 240

Qy 422 CCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATC 481
Db 241 CCCAGACCTTTCACCCACGTGAACACCACCACCAGCGCCACACATAGCACAGCCACCATC 300

Qy 482 CTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGG 541
Db 301 CTCAACCCTCGAGATACGTACTGCAGGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGG 360

Qy 542 GACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCC 601
Db 361 GACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATGTCTTCC 420

Qy 602 CCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTG 661
Db 421 CCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACCTACCTG 480

Qy 662 GTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCC 721
Db 481 GTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATCCACCCC 540

Qy 722 AGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTC 781
Db 541 AGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTGATCTTC 600

Qy 782 ACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAAC 841
Db 601 ACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATCCTAAAC 660

Qy 842 ACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTG 901
Db 661 ACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTG 720

Qy 902 AGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTT 961
Db 721 AGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTT 780

Qy 962 TCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATT 1021
Db 781 TCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATT 840

Qy 1022 ATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACA 1061
Db 841 ATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACA 880

RESULT 4
AAS76631
ID AAS76631 standard; cDNA; 1137 BP.
XX
AC AAS76631;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12435.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX

```
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12444.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 12435; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1137 BP; 332 A; 259 C; 254 G; 292 T; 0 other;

  Query Match             42.9%;  Score 694;  DB 23;  Length 1137;
  Best Local Similarity   99.5%;  Pred. No. 4.6e-294;
  Matches  944;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     536 GTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGATG 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      43 GTGAGGGACCACTTGGGACGCAGGAAGCAATATGGCGGGGATTTCCTGAGGGCCAGGAGG 102

Qy     596 TCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACC 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     103 TCTTCCCCAGCGCTGATGGCAGGTGCTTCAGGAAAGGTGACTGACTTCAACAACGGCACC 162

Qy     656 TACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATC 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     163 TACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCTGCTGCTCATC 222

Qy     716 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     223 CACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTATGACAGGGTG 282

Qy     776 ATCTTCACTGGCCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATC 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     283 ATCTTCACTGGCCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATGTGGCCTGATC 342

Qy     836 CTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGAAGGCTTCTAC 895
           |||||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db     343 CTAAACACAAATGCTGAATTGTGCCAATACCTGGACAACAGAGACCAAGAAAGCTTCTAC 402

Qy     896 TGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAG 955
           || |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     403 TGGGTGAGGCCTCAACACATGCGCTGTGCTGCACTCACTCACATGTATTCTAAGAACAAG 462

Qy     956 AAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     463 AAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAATGTGGGTGTA 522

Qy    1016 GAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     523 GAGATTATGGAAAAATTCAATACAATTAGTGTCTCCAAATGCAACACACTGAAGTCAGTG 582

Qy    1076 GATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATC 1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     583 GATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATTTGGATAGGAACATC 642

Qy    1136 AACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAA 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     643 AACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGACCTATTCAGTCAAA 702

Qy    1196 GAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAAATACTGTCATT 1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     703 GAGATGGAGTACCTCACCCGGGCCATTGACAGAACTGGAGGAGAAAAAAATACTGTCATT 762

Qy    1256 GTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCC 1315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     763 GTTATTTCCCTGGGCCAGCATTTCAGACCCTTTCCCATTGATGTTTTTATCCGAAGGGCC 822

Qy    1316 CTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATC 1375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     823 CTCAATGTCCACAAAGCCATTCAGCATCTTCTTCTGAGAAGCCCAGACACTATGGTTATC 882

Qy    1376 ATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCAT 1435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     883 ATCAAAACAGAAAACATCAGGGAGATGTACAATGATGCAGAAAGATTTAGTGACTTTCAT 942

Qy    1436 GGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTG 1484
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     943 GGTTACATTCAATATCTCATCATAAAGGACATTTTCCAGGATCTCAGTG 991


RESULT 5
ABN61431
ID   ABN61431 standard; cDNA; 400 BP.
XX
AC   ABN61431;
XX
DT   28-JUN-2002  (first entry)
XX
DE   Human cancer related polynucleotide SEQ ID NO 1398.
XX
KW   Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW   gene therapy; cancer; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200214500-A2.
XX
PD   21-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-US25840.
XX
PR   16-AUG-2000; 2000US-226326P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Sudduth-Klinger J,  Reinhard C,  Randazzo F;
PI   Lamson G,  Scott EM,  Zhang G,  Kassam A,  Pot D,  Labat I;
XX
DR   WPI; 2002-241905/29.
XX
PT   New nucleic acid for producing a polypeptide, detecting differentially
PT   expressed genes correlated with a cancerous state of a mammalian cell,
PT   and inhibiting tumor growth -
XX
PS   Claim 1; SEQ ID NO 1398; 883pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC   with cytostatic activity. The polynucleotide is used to produce a
CC   polypeptide, to detect differentially expressed genes correlated with a
CC   cancerous state of a mammalian cell and to inhibit tumour growth. The
```

```
CC   polynucleotide is used as a probe in mapping and tissue profiling. The
CC   encoded polypeptide and antibodies to the polypeptide can also be used
CC   for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC   gene therapy.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 400 BP; 110 A; 98 C; 92 G; 100 T; 0 other;

  Query Match             24.8%;  Score 400;  DB 24;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-165;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     646 CAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCT 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCTCTGTCTCT 60

Qy     706 GCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTA 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCTGCTCATCCACCCCAGTGAAGGGGTGTCAGCTCTCTGGAGTGCAAGGAACCAAGGCTA 120

Qy     766 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGACAGGGTGATCTTCACTGGCCAGTTTGTCAATGGCACTTCCCAAGTCCACTCTGAATG 180

Qy     826 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAGACCAAGA 240

Qy     886 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTGCTGCACTCACTCACATGTATTC 300

Qy     946 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAA 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAAAGAGCCTCTTTGAAAGGTCAAA 360

Qy    1006 TGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTAGT 1045
           |||||||||||||||||||||||||||||||||||||||
Db     361 TGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTAGT 400


RESULT 6
AAS92089
ID   AAS92089 standard; cDNA; 437 BP.
XX
AC   AAS92089;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #27893.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG27902.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 27893; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 437 BP; 150 A; 101 C; 70 G; 116 T; 0 other;

  Query Match             23.9%;  Score 386;  DB 23;  Length 437;
  Best Local Similarity   99.8%;  Pred. No. 4.2e-159;
  Matches  436;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 AGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGTATGCCA 71
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AGGTAAACAAACTGCAACTTATATCTGCAATTTATTTTGGTATAGACAAGAGGGATGCCA 60

Qy      72 GTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGT 131
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTAGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGT 120

Qy     132 ATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATT 191
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATGGGATTGAACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATT 180

Qy     192 ATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCC 251
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATAAGTCACTATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCC 240

Qy     252 AGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACA 311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGAACTCCACAAAGGTTTGGTCTGCTCTAAACTTATCCATCTCCCTCCATTACTGGAACA 300

Qy     312 ACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGA 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACTCCACAAAGTCCTTATTCCCTAAAACACCACTGATATCATTAAAGCCACTAACAGAGA 360

Qy     372 CTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTT 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTT 420

Qy     432 TCACCCACGTGAACACC 448
           |||||||||||||||||
Db     421 TCACCCACGTGAACACC 437


RESULT 7
ABL36592
ID   ABL36592 standard; cDNA; 557 BP.
XX
```

```
AC   ABL36592;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:181.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US18557.
XX
PR   09-JUN-2000; 2000US-210899P.
PR   20-FEB-2001; 2001US-270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Harlocker SL, Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient -
XX
PS   Claim 1; SEQ ID 181; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour.
XX
SQ   Sequence 557 BP; 192 A; 112 C; 121 C; 130 T; 2 other;

  Query Match             12.3%;  Score 198;  DB 24;  Length 557;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-77;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     864 ACCTGGACAACAGACACCAACAAGCCTTCTACTGTGTCACGCCTCAACACATCCCCTGTG 923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACCTGGACAACAGAGACCAAGAAGGCTTCTACTGTGTGAGGCCTCAACACATGCCCTGTG 60

Qy     924 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGCACTCACTCACATGTATTCTAAGAACAAGAAAGTTTCTTATCTTAGCAAACAAGAAA 120

Qy     984 AGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTA 1043
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAGCCTCTTTGAAAGGTCAAATGTGGGTGTAGAGATTATGGAAAAATTCAATACAATTA 180

Qy    1044 GTGTCTCCAAATCCAACA 1061
           ||||||||||||||||||
Db     181 GTGTCTCCAAATGCAACA 198


RESULT 8
AAF93741
ID   AAF93741 standard; cDNA; 275 BP.
XX
AC   AAF93741;
XX
DT   21-MAY-2001  (first entry)
XX
DE   cDNA encoding SRT protein isolated from testis tissue SEQ ID 562.
XX
KW   Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX
OS   Homo sapiens.
XX
PN   WO200107611-A2.
XX
PD   01-FEB-2001.
XX
PF   21-JUL-2000; 2000WO-US20006.
XX
PR   26-JUL-1999;  99US-0145701.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP, Goddard A, Wood WI;
XX
DR   WPI; 2001-112729/12.
XX
PT   New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT   for production of recombinant SRT polypeptides, gene mapping,
PT   diagnosing genetic disorders and for gene therapy -
XX
PS   Claim 2; Fig 562; 663pp; English.
XX
CC   Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding
CC   human SRT proteins. The cDNA sequences are isolated from various
CC   different human tissue cDNA libraries. The invention relates to a method
CC   for detecting cDNA encoding an SRT protein, a vector containing cDNA
CC   encoding SRT, a host cell transformed with the vector, an isolated SRT
CC   polypeptide, and an antibody which binds to SRT. The polynucleotide
CC   sequence can be used in gene therapy and is useful in the recombinant
CC   production of SRT polypeptides, as a hybridisation probe to screen
CC   libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC   map the gene encoding the SRT polypeptides and analysing genetic
CC   disorders, tissue typing and disease tissue detection. The SRT
CC   polynucleotide sequences can be used in polymerase chain reaction,
CC   screening for new therapeutic molecules and generation of antisense RNA
CC   and DNA.
XX
SQ   Sequence 275 BP; 83 A; 65 C; 43 G; 81 T; 3 other;

  Query Match             11.9%;  Score 192;  DB 22;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-74;
  Matches  192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      81 TGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGATTG 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14 TGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTCGCCAGAGAGTCTATGTATGGGATTG 73

Qy     141 AACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTCAC 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 AACAATCTGTAAACTAAAGGATCCTAATCATGAAAATAAGTATGATAAATTATAAGTCAC 133

Qy     201 TATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCA 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 TATTGGCACTGTTGTTTATATTAGCCTCCTGGATCATTTTTACAGTTTTCCAGAACTCCA 193

Qy     261 CAAAGGTTTGGT 272
           ||||||||||||
Db     194 CAAAGGTTTCCT 205


RESULT 9
AAK88279
ID   AAK88279 standard; cDNA; 817 BP.
XX
AC   AAK88279;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human digestive system antigen coding sequence SEQ ID NO: 595.
```

XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

```
RESULT 11
ABL36799/c
ID   ABL36799 standard; cDNA; 583 BP.
XX
AC   ABL36799;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:388.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US18557.
XX
PR   09-JUN-2000; 2000US-210899P.
PR   20-FEB-2001; 2001US-270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Harlocker SL,  Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient -
XX
PS   Claim 1; SEQ ID 388; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour.
XX
SQ   Sequence 583 BP; 186 A; 110 C; 126 G; 157 T; 4 other;

  Query Match             5.8%;  Score 93;  DB 24;  Length 583;
  Best Local Similarity   99.3%;  Pred. No. 9.7e-31;
  Matches  143;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1062 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 1121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 CACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACCAGCTTGCTGTGGATT 85

Qy   1122 TGGATAGGAACATCAACATCCAGTGGCAAAAATATTGTTATCCCTTGATAGGATCAATGA 1181
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     84 TGGATAGGAACATCAACATCCAGTGGCAAAAACATTGTTATCCCTTGATAGGATCAATGA 25

Qy   1182 CCTATTCAGTCAAAGAGATGGAGT 1205
           ||||||||||||||||||||||||
Db     24 CCTATTCAGTCAAAGAGATGGAGT 1

RESULT 12
ABN60352
ID   ABN60352 standard; cDNA; 211 BP.
XX
AC   ABN60352;
XX
DT   28-JUN-2002  (first entry)
XX
DE   Human cancer related polynucleotide SEQ ID NO 319.
XX
KW   Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW   gene therapy; cancer; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200214500-A2.
XX
PD   21-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-US25840.
XX
PR   16-AUG-2000; 2000US-226326P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Sudduth-Klinger J,  Reinhard C,  Randazzo F;
PI   Lamson G,  Scott EM,  Zhang G,  Kassam A,  Pot D,  Labat I;
XX
DR   WPI; 2002-241905/29.
XX
PT   New nucleic acid for producing a polypeptide, detecting differentially
PT   expressed genes correlated with a cancerous state of a mammalian cell,
PT   and inhibiting tumor growth -
XX
PS   Claim 1; SEQ ID NO 319; 883pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC   with cytostatic activity. The polynucleotide is used to produce a
CC   polypeptide, to detect differentially expressed genes correlated with a
CC   cancerous state of a mammalian cell and to inhibit tumour growth. The
CC   polynucleotide is used as a probe in mapping and tissue profiling. The
CC   encoded polypeptide and antibodies to the polypeptide can also be used
CC   for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC   gene therapy.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 211 BP; 75 A; 68 C; 32 G; 36 T; 0 other;

  Query Match             5.4%;  Score 88;  DB 24;  Length 211;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-28;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     31 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 90

Qy    411 AGCAGATCCCACCCAGACCTTTCACCCA 438
          ||||||||||||||||||||||||||||
Db     91 AGCAGATCCCACCCAGACCTTTCACCCA 118

RESULT 13
ABN60393
ID   ABN60393 standard; cDNA; 611 BP.
XX
AC   ABN60393;
XX
DT   28-JUN-2002  (first entry)
XX
DE   Human cancer related polynucleotide SEQ ID NO 360.
XX
KW   Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW   gene therapy; cancer; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200214500-A2.
```

XX
PD 21-FEB-2002.
XX
PF 16-AUG-2001; 2001WO-US25840.
XX
PR 16-AUG-2000; 2000US-226326P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
DR WPI; 2002-241905/29.
XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth -
XX
PS Claim 1; SEQ ID NO 360; 883pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 611 BP; 188 A; 163 C; 119 G; 141 T; 0 other;

Query Match 5.4%; Score 88; DB 24; Length 611;
Best Local Similarity 100.0%; Pred. No. 1.5e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 305 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 364

Qy 411 AGCAGATCCCACCCAGACCTTTCACCCA 438
||||||||||||||||||||||||||||
Db 365 AGCAGATCCCACCCAGACCTTTCACCCA 392

RESULT 14
ABK69141
ID ABK69141 standard; cDNA; 684 BP.
XX
AC ABK69141;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human secreted protein, SEQ ID No 65.
XX
KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200224721-A1.
XX
PD 28-MAR-2002.
XX
PF 09-JAN-2001; 2001WO-US00544.
XX
PR 20-SEP-2000; 2000US-234211P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Baker KP, Rosen CA, Birse CE, Soppet DR, Olsen HS;
PI Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR, Shi Y;
PI Choi GH;
XX
DR WPI; 2002-330012/36.
DR P-PSDB; AAU96220.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; Page 491; 562pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) encoding a
CC human secreted protein (II). (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. ABK69078-ABK69143
CC represent human secreted protein coding sequences, PCR primers and
CC related sequences used in cloning and expression of the secreted
CC proteins described in examples of the invention.
XX
SQ Sequence 684 BP; 189 A; 185 C; 147 G; 158 T; 5 other;

Query Match 5.4%; Score 88; DB 24; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.5e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 294 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 353

Qy 411 AGCAGATCCCACCCAGACCTTTCACCCA 438
||||||||||||||||||||||||||||
Db 354 AGCAGATCCCACCCAGACCTTTCACCCA 381

RESULT 15
ABK69140
ID ABK69140 standard; cDNA; 1177 BP.
XX
AC ABK69140;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human secreted protein, SEQ ID No 64.
XX
KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.

XX
OS Homo sapiens.
XX
PN WO200224721-A1.
XX
PD 28-MAR-2002.
XX
PF 09-JAN-2001; 2001WO-US00544.
XX
PR 20-SEP-2000; 2000US-234211P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Baker KP, Rosen CA, Birse CE, Soppet DR, Olsen HS;
PI Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR, Shi Y;
PI Choi GH;
XX
DR WPI; 2002-330012/36.
DR P-PSDB; AAU96219.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; Page 490; 562pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) encoding a
CC human secreted protein (II). (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. ABK69078-ABK69143
CC represent human secreted protein coding sequences, PCR primers and
CC related sequences used in cloning and expression of the secreted
CC proteins described in examples of the invention.
XX
SQ Sequence 1177 BP; 347 A; 289 C; 253 G; 273 T; 15 other;

Query Match 5.4%; Score 88; DB 24; Length 1177;
Best Local Similarity 100.0%; Pred. No. 1.4e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     351 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 CATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAAACTAGATC 416

Qy     411 AGCAGATCCCACCCAGACCTTTCACCCA 438
           ||||||||||||||||||||||||||||
Db     417 AGCAGATCCCACCCAGACCTTTCACCCA 444
```

Search completed: November 22, 2002, 02:43:20
Job time : 340 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:35:41 ; Search time 83 Seconds
(without alignments)
5970.960 Million cell updates/sec

Title: US-09-729-454-3
Perfect score: 1616
Sequence: 1 atccatgctaaaggtaaaca.....aataacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	22	1.4	607	4	US-09-385-982-288	Sequence 288, App
	2	20	1.2	1967	4	US-09-398-395A-23	Sequence 23, Appl
	3	20	1.2	1968	2	US-08-937-540-3	Sequence 3, Appli
	4	20	1.2	29604	3	US-08-781-891-207	Sequence 207, App
c	5	19	1.2	259	3	US-09-103-359-13	Sequence 13, Appl
	6	19	1.2	425	4	US-09-247-155-73	Sequence 73, Appl
c	7	19	1.2	485	4	US-09-385-982-31	Sequence 31, Appl
c	8	19	1.2	1005	3	US-09-103-359-4	Sequence 4, Appli
	9	19	1.2	1356	1	US-08-426-169-2	Sequence 2, Appli
	10	19	1.2	1356	4	US-09-233-813-2	Sequence 2, Appli
	11	19	1.2	1356	5	PCT-US95-09470-2	Sequence 2, Appli
	12	19	1.2	1615	4	US-08-308-814-1	Sequence 1, Appli
	13	19	1.2	1809	1	US-08-455-001-1	Sequence 1, Appli
	14	19	1.2	1809	5	PCT-US95-11869-1	Sequence 1, Appli
c	15	19	1.2	2038	4	US-09-276-531-37	Sequence 37, Appl
	16	19	1.2	2304	1	US-08-464-266-1	Sequence 1, Appli
	17	19	1.2	2304	1	US-08-464-272-1	Sequence 1, Appli
	18	19	1.2	2304	4	US-08-464-514-1	Sequence 1, Appli
	19	19	1.2	2304	4	US-08-486-403-1	Sequence 1, Appli
	20	19	1.2	2981	1	US-08-257-073-2	Sequence 2, Appli
	21	19	1.2	2981	2	US-08-184-009-119	Sequence 119, App
	22	19	1.2	2981	2	US-08-458-356-119	Sequence 119, App
	23	19	1.2	2981	4	US-08-460-736-119	Sequence 119, App
	24	19	1.2	3107	4	US-08-213-419B-1	Sequence 1, Appli
	25	19	1.2	6124	4	US-08-213-419B-3	Sequence 3, Appli
c	26	18	1.1	45	4	US-09-025-639-1	Sequence 1, Appli
c	27	18	1.1	45	4	US-09-333-237-1	Sequence 1, Appli
c	28	18	1.1	60	4	US-09-410-960-1	Sequence 1, Appli
c	29	18	1.1	60	4	US-09-410-960-3	Sequence 3, Appli
c	30	18	1.1	60	4	US-09-410-960-4	Sequence 4, Appli
	31	18	1.1	269	4	US-09-149-476-142	Sequence 142, App
c	32	18	1.1	567	4	US-09-385-982-427	Sequence 427, App
	33	18	1.1	730	4	US-09-352-990-3	Sequence 3, Appli
	34	18	1.1	809	1	US-08-441-629-1	Sequence 1, Appli
	35	18	1.1	809	3	US-08-776-207-1	Sequence 1, Appli
	36	18	1.1	809	4	US-09-507-773-1	Sequence 1, Appli
	37	18	1.1	809	5	PCT-US95-09172-1	Sequence 1, Appli
	38	18	1.1	815	4	US-09-288-143-24	Sequence 24, Appl
	39	18	1.1	897	4	US-08-899-330-14	Sequence 14, Appl
	40	18	1.1	904	4	US-09-171-209-24	Sequence 24, Appl
	41	18	1.1	1356	1	US-07-603-133B-1	Sequence 1, Appli
	42	18	1.1	1356	1	US-08-426-169-3	Sequence 3, Appli
	43	18	1.1	1356	4	US-09-233-813-3	Sequence 3, Appli
	44	18	1.1	1356	5	PCT-US95-09470-3	Sequence 3, Appli
	45	18	1.1	1460	4	US-09-370-838-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1
US-09-385-982-288/c
; Sequence 288, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 288
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(607)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-288

Query Match 1.4%; Score 22; DB 4; Length 607;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1595 TAAATAACAAAAAAAAAAAAAA 1616
||||||||||||||||||||||
Db 35 TAAATAACAAAAAAAAAAAAAA 14

RESULT 2
US-09-398-395A-23
; Sequence 23, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A

```
;  CURRENT FILING DATE:  1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;    LENGTH: 1967
;    TYPE: DNA
;    ORGANISM: Salvia officinalis
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (13)...(1785)
;    OTHER INFORMATION: 1,8-cineole synthase
US-09-398-395A-23

  Query Match             1.2%;  Score 20;  DB 4;  Length 1967;
  Best Local Similarity   100.0%;  Pred. No. 9.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1597 AATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||||
Db   1939 AATAACAAAAAAAAAAAAAA 1958


RESULT 3
US-08-937-540-3
; Sequence 3, Application US/08937540
; Patent No. 5891697
;  GENERAL INFORMATION:
;    APPLICANT:  Croteau, Rodney B
;    APPLICANT:  Wise, Mitchell L
;    APPLICANT:  Savage, Thomas J
;    APPLICANT:  Katahira, Eva J
;    TITLE OF INVENTION:  Monoterpene Synthases from Common Sage
;    TITLE OF INVENTION:  (Salvia officinalis)
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
;      STREET:  1420 FIFTH AVENUE
;      CITY:  SEATTLE
;      STATE:  WASHINGTON
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/937,540
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K
;      REGISTRATION NUMBER:  26,997
;      REFERENCE/DOCKET NUMBER:  WSUR111254
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206 695 1718
;      TELEFAX:  206 224 0779
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1968 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Salvia officinalis
;    IMMEDIATE SOURCE:
;      CLONE:  1,8-cineole synthase
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14..1788
US-08-937-540-3

  Query Match             1.2%;  Score 20;  DB 2;  Length 1968;
  Best Local Similarity   100.0%;  Pred. No. 9.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1597 AATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||||
Db   1940 AATAACAAAAAAAAAAAAAA 1959


RESULT 4
US-08-781-891-207
; Sequence 207, Application US/08781891
; Patent No. 6090620
;  GENERAL INFORMATION:
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Yu, Chang-En
;    APPLICANT:  Oshima, Junko
;    APPLICANT:  Mulligan, John T.
;    APPLICANT:  Schellenberg, Gerald D.
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
;    TITLE OF INVENTION:  WERNER'S SYNDROME
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,891
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6090620tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  240052.419
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  207:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29604 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-207

  Query Match             1.2%;  Score 20;  DB 3;  Length 29604;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1597 AATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||||
Db  12364 AATAACAAAAAAAAAAAAAA 12383


RESULT 5
US-09-103-359-13/c
; Sequence 13, Application US/09103359
```

```
, Patent No. 6057108
   GENERAL INFORMATION:
,    APPLICANT:  Hillman, Jennifer L.
,    APPLICANT:  Bandman, Olga
     APPLICANT:  Guegler, Karl J.
,    APPLICANT:  Corley, Neil C.
     APPLICANT:  Yue, Henry
,    APPLICANT:  Patterson, Chandra
,    TITLE OF INVENTION:  HUMAN ARF-RELATED PROTEINS
     NUMBER OF SEQUENCES:  14
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
,      STREET:  3174 Porter Drive
,      CITY:  Palo Alto
,      STATE:  CA
,      COUNTRY:  USA
       ZIP:  94304
,    COMPUTER READABLE FORM:
,      MEDIUM TYPE:  Diskette
       COMPUTER:  IBM Compatible
       OPERATING SYSTEM:  Windows
,      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
     CURRENT APPLICATION DATA:
,      APPLICATION NUMBER:  US/09/103,359
,      FILING DATE:  HEREWITH
,      CLASSIFICATION:
,    PRIOR APPLICATION DATA:
       APPLICATION NUMBER:
       FILING DATE:
     ATTORNEY/AGENT INFORMATION:
       NAME:  Cerrone, Michael C
,      REGISTRATION NUMBER:  39,132
       REFERENCE/DOCKET NUMBER:  PF-0537 US
,    TELECOMMUNICATION INFORMATION:
       TELEPHONE:  650-855-0555
,      TELEFAX:  650-855-0572
       TELEX:
   INFORMATION FOR SEQ ID NO:  13:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  259 base pairs
,      TYPE:  nucleic acid
,      STRANDEDNESS:  single
,      TOPOLOGY:  linear
     IMMEDIATE SOURCE:
,      LIBRARY:  GBLANOT01
       CLONE:  2614213H1
US-09-103-359-13

  Query Match             1.2%;  Score 19;  DB 3;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     939 TGTATTCTAAGAACAAGAA 957
           |||||||||||||||||||
Db      34 TGTATTCTAAGAACAAGAA 16


RESULT 6
US-09-247-155-73
, Sequence 73, Application US/09247155A
, Patent No. 6312922
, GENERAL INFORMATION:
   APPLICANT: Dumas Milne Edwards, Jean-Baptiste
,  APPLICANT: Duclert, Aymeric
   APPLICANT: Bouqueleret, Lydie
,  TITLE OF INVENTION: Complementary DNAs
,  FILE REFERENCE: GENSET.021A
   CURRENT APPLICATION NUMBER: US/09/247,155A
   CURRENT FILING DATE: 1999-02-09
,  EARLIER APPLICATION NUMBER: 60/074,121
   EARLIER FILING DATE: 1998-02-09
,  EARLIER APPLICATION NUMBER: 60/081,563
   EARLIER FILING DATE: 1998-04-13
;  EARLIER APPLICATION NUMBER: 60/096,116
;  EARLIER FILING DATE: 1998-08-10
;  EARLIER APPLICATION NUMBER: 60/099,273
;  EARLIER FILING DATE: 1998-10-04
;  NUMBER OF SEQ ID NOS: 182
;  SOFTWARE: Patent.pm
; SEQ ID NO 73
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 55..291
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 55..255
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 4.4
;   OTHER INFORMATION: seq LISLVASLFMGFG/VL
;   FEATURE:
;   NAME/KEY: polyA_signal
;   LOCATION: 390..395
;   FEATURE:
;   NAME/KEY: polyA_site
;   LOCATION: 410..425
US-09-247-155-73

  Query Match             1.2%;  Score 19;  DB 4;  Length 425;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1598 ATAACAAAAAAAAAAAAAA 1616
           ||||||||||||||||||||
Db     405 ATAACAAAAAAAAAAAAAA 423


RESULT 7
US-09-385-982-31/c
; Sequence 31, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 31
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(485)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-31

  Query Match             1.2%;  Score 19;  DB 4;  Length 485;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     939 TGTATTCTAAGAACAAGAA 957
           |||||||||||||||||||
Db      54 TGTATTCTAAGAACAAGAA 36
```

```
RESULT 8
US-09-103-359-4/c
; Sequence 4, Application US/09103359
; Patent No. 6057108
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Patterson, Chandra
;    TITLE OF INVENTION:  HUMAN ARF-RELATED PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/103,359
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0537 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-855-0572
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1005 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  COLNNOT13
;      CLONE:  1333754
JS-09-103-359-4

  Query Match             1.2%;  Score 19;  DB 3;  Length 1005;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    939 TGTATTCTAAGAACAAGAA 957
          |||||||||||||||||||
Db     33 TGTATTCTAAGAACAAGAA 15


RESULT 9
JS-08-426-169-2
  Sequence 2, Application US/08426169
  Patent No. 5620896
  GENERAL INFORMATION:
    APPLICANT:  Herrmann, John E.
    APPLICANT:  Robinson, Harriet L.
    APPLICANT:  Fynan, Ellen F.
    TITLE OF INVENTION:  DNA VACCINES AGAINST ROTAVIRUS
    TITLE OF INVENTION:  INFECTIONS
    NUMBER OF SEQUENCES:  6
    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/426,169
;      FILING DATE:  20-APR-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04020/049001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1356 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Human rotavirus VP6
US-08-426-169-2

  Query Match             1.2%;  Score 19;  DB 1;  Length 1356;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1408 TGATGCAGAAAGATTTAGT 1426
          |||||||||||||||||||
Db    704 TGATGCAGAAAGATTTAGT 722


RESULT 10
US-09-233-813-2
; Sequence 2, Application US/09233813
; Patent No. 6165993
; GENERAL INFORMATION:
;  APPLICANT: Herrmann, John E.
;  APPLICANT: Robinson, Harriet L.
;  APPLICANT: Fynan, Ellen F.
;  TITLE OF INVENTION: DNA VACCINES AGAINST ROTAVIRUS INFECTIONS
;  FILE REFERENCE: 04020/049003
;  CURRENT APPLICATION NUMBER: US/09/233,813
;  CURRENT FILING DATE: 1998-03-02
;  EARLIER APPLICATION NUMBER: US 08/842,563
;  EARLIER FILING DATE: 1997-04-15
;  EARLIER APPLICATION NUMBER: US 08/426,169
;  EARLIER FILING DATE: 1995-04-20
;  EARLIER APPLICATION NUMBER: US 08/187,879
;  EARLIER FILING DATE: 1994-01-27
;  EARLIER APPLICATION NUMBER: US 08/009,833
;  EARLIER FILING DATE: 1993-01-27
;  EARLIER APPLICATION NUMBER: US 07/855,562
;  EARLIER FILING DATE: 1992-03-23
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1356
;   TYPE: DNA
;   ORGANISM: Human rotavirus VP6
US-09-233-813-2

  Query Match             1.2%;  Score 19;  DB 4;  Length 1356;
```

```
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1408 TGATGCAGAAAGATTTAGT 1426
          ||||||||||||||||||
Db    704 TGATGCAGAAAGATTTAGT 722


RESULT 11
PCT-US95-09470-2
, Sequence 2, Application PC/TUS9509470
,  GENERAL INFORMATION:
,    APPLICANT:  UNIVERSITY OF MASSACHUSETTS MEDICAL CENTER
,    TITLE OF INVENTION:  DNA VACCINES AGAINST ROTAVIRUS
,    TITLE OF INVENTION:  INFECTIONS
,    NUMBER OF SEQUENCES:  6
,    CORRESPONDENCE ADDRESS:
,      ADDRESSEE:  Fish & Richardson P.C.
,      STREET:  225 Franklin Street
,      CITY:  Boston
       STATE:  MA
       COUNTRY:  USA
       ZIP:  02110-2804
,    COMPUTER READABLE FORM:
,      MEDIUM TYPE:  Floppy disk
,      COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
,      SOFTWARE:  PatentIn Release #1.0, Version #1.25
,    CURRENT APPLICATION DATA:
,      APPLICATION NUMBER:  PCT/US95/09470
,      FILING DATE:  26-JUL-95
,      CLASSIFICATION:
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US 08/426,169
,      FILING DATE:  20-APR-1995
,      CLASSIFICATION:
     ATTORNEY/AGENT INFORMATION:
       NAME:  Fasse, J. Peter
,      REGISTRATION NUMBER:  32,983
,      REFERENCE/DOCKET NUMBER:  04020/049W01
,    TELECOMMUNICATION INFORMATION:
,      TELEPHONE:  (617)542-5070
,      TELEFAX:  (617) 542-8906
   INFORMATION FOR SEQ ID NO:  2:
,    SEQUENCE CHARACTERISTICS:
,      LENGTH:  1356 base pairs
       TYPE:  nucleic acid
,      STRANDEDNESS:  single
       TOPOLOGY:  linear
,    MOLECULE TYPE:  cDNA
     ORIGINAL SOURCE:
       ORGANISM:  Human rotavirus VP6
PCT-US95-09470-2

  Query Match             1.2%;  Score 19;  DB 5;  Length 1356;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1408 TGATGCAGAAAGATTTAGT 1426
          ||||||||||||||||||
Db    704 TGATGCAGAAAGATTTAGT 722


RESULT 12
US-08-308-814-1
, Sequence 1, Application US/08308814
, Patent No. 6268476
,  GENERAL INFORMATION:
,    APPLICANT:  Flanagan, John G.
,    APPLICANT:  Cheng, Hwai-Jong
,    TITLE OF INVENTION:  EPH Receptor Ligands, and Uses Related
,    TITLE OF INVENTION:  Thereto
,    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(txt)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/308,814
;      FILING DATE:  19-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36.709
;      REFERENCE/DOCKET NUMBER:  HMI-011
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1615 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  10..636
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  10..69
US-08-308-814-1

  Query Match             1.2%;  Score 19;  DB 4;  Length 1615;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1598 ATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||
Db   1534 ATAACAAAAAAAAAAAAAA 1552


RESULT 13
US-08-455-001-1
; Sequence 1, Application US/08455001
; Patent No. 5795734
;  GENERAL INFORMATION:
;    APPLICANT:  Flanagan, John G.
;    APPLICANT:  Cheng, Hwai-Jong
;    TITLE OF INVENTION:  EPH Receptor Ligands, and Uses Related
;    TITLE OF INVENTION:  Thereto
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII (text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,001
;      FILING DATE:  31 MAY 1995
;      CLASSIFICATION:  800
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  HMI-011CP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1809 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  265..891
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  892..1809
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  255..324
US-08-455-001-1

  Query Match             1.2%;  Score 19;  DB 1;  Length 1809;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1598 ATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||
Db   1789 ATAACAAAAAAAAAAAAAA 1807


RESULT 14
PCT-US95-11869-1
; Sequence 1, Application PC/TUS9511869
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  EPH Receptor Ligands, and Uses Related
;    TITLE OF INVENTION:  Thereto
;    NUMBER OF SEQUENCES:  5
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/11869
;      FILING DATE:  19-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  HMI-011CPPC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1809 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  265..891
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  892..1809
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  255..324
PCT-US95-11869-1

  Query Match             1.2%;  Score 19;  DB 5;  Length 1809;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1598 ATAACAAAAAAAAAAAAAA 1616
          ||||||||||||||||||
Db   1789 ATAACAAAAAAAAAAAAAA 1807


RESULT 15
US-09-276-531-37/c
; Sequence 37, Application US/09276531
; Patent No. 6183968
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Reddy, Roopa
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF GENES ENCODING
;    TITLE OF INVENTION:  RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATI N
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/276,531
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/079,677
;      FILING DATE:  March 27, 1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lynn E. Murry, Ph.D.
;      REGISTRATION NUMBER:  42,918
;      REFERENCE/DOCKET NUMBER:  PA-0008 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2038 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSTUT05
;      CLONE:  838871
US-09-276-531-37

  Query Match             1.2%;  Score 19;  DB 4;  Length 2038;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      353 TTAAAGCCACTAACAGAGA 371
            ||||||||||||||||||||
Db     1071 TTAAAGCCACTAACAGAGA 1053
```

Search completed: November 22, 2002, 04:31:19
Job time : 128 secs

US 097294540DP1

Creation date: 12-01-2003
Indexing Officer: IMULAMBA - ISIDORE MULAMBA
Team: OIPEBackFileIndexing
Dossier: 09729454

Legal Date: 11-27-2002

No.	Doccode	Number of pages
1	CTNF	9
2	892	1
3	NPL	3
4	NPL	8
5	NPL	2
6	NPL	4

Total number of pages: 27

Remarks:

Order of re-scan issued on .....